

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 31-03-2015		2. REPORT TYPE Ph.D. Dissertation		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE High-Order Accurate Solutions to the Helmholtz Equation in the Presence of Boundary Singularities			5a. CONTRACT NUMBER W911NF-11-1-0384		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611102		
6. AUTHORS Darrel Steven Britt			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES North Carolina State University 2701 Sullivan Drive Raleigh, NC 27695 -7514			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 59637-MA.20		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT Problems of time-harmonic wave propagation arise in important fields of study such as geological surveying, radar detection/evasion, and aircraft design. These often involve high-frequency waves, which demand high-order methods to mitigate the dispersion error. We propose a high-order method for computing solutions to the variable-coefficient inhomogeneous Helmholtz equation in two dimensions on domains bounded by piecewise smooth curves of arbitrary shape with a finite number of boundary singularities at known locations.					
15. SUBJECT TERMS difference potentials, boundary projections, Calderon's operators, regular grids, curvilinear boundaries, non-conforming boundaries, solutions with singularities, singularity subtraction, discontinuous boundary data, asymptotic expansion near singularity, compact					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU	UU		Semyon Tsynkov
					19b. TELEPHONE NUMBER 919-515-1877

Report Title

High-Order Accurate Solutions to the Helmholtz Equation in the Presence of Boundary Singularities

ABSTRACT

Problems of time-harmonic wave propagation arise in important fields of study such as geological surveying, radar detection/evasion, and aircraft design. These often involve high-frequency waves, which demand high-order methods to mitigate the dispersion error. We propose a high-order method for computing solutions to the variable-coefficient inhomogeneous Helmholtz equation in two dimensions on domains bounded by piecewise smooth curves of arbitrary shape with a finite number of boundary singularities at known locations.

We utilize compact finite difference (FD) schemes on regular structured grids to achieve high-order accuracy due to their efficiency and simplicity, as well as the capability to approximate variable-coefficient differential operators. In this work, a 4th-order compact FD scheme for the variable-coefficient Helmholtz equation on a Cartesian grid in 2D is derived and tested. The well known limitation of finite differences is that they lose accuracy when the boundary curve does not coincide with the discretization grid, which is a severe restriction on the geometry of the computational domain. Therefore, the algorithm presented in this work combines high-order FD schemes with the method of difference potentials (DP), which retains the efficiency of FD while allowing for boundary shapes that are not aligned with the grid without sacrificing the accuracy of the FD scheme.

Additionally, the theory of DP allows for the universal treatment of the boundary conditions. One of the significant contributions of this work is the development of an implementation that accommodates general boundary conditions (BCs). In particular, Robin BCs with discontinuous coefficients are studied, for which we introduce a piecewise parameterization of the boundary curve. Problems with discontinuities in the boundary data itself are also studied.

We observe that the design convergence rate suffers whenever the solution loses regularity due to the boundary conditions. This is because the FD scheme is only consistent for classical solutions of the PDE. For this reason, we implement the method of singularity subtraction as a means for restoring the design accuracy of the scheme in the presence of singularities at the boundary. While this method is well studied for low order methods and for problems in which singularities arise from the geometry (e.g., corners), we adapt it to our high-order scheme for curved boundaries via a conformal mapping and show that it can also be used to restore accuracy when the singularity arises from the BCs rather than the geometry.

Altogether, the proposed methodology for 2D boundary value problems is computationally efficient, easily handles a wide class of boundary conditions and boundary shapes that are not aligned with the discretization grid, and requires little modification for solving new problems.

ABSTRACT

BRITT, JR., DARRELL STEVEN. High-Order Accurate Solutions to the Helmholtz Equation in the Presence of Boundary Singularities. (Under the direction of Semyon Tsynkov.)

Problems of time-harmonic wave propagation arise in important fields of study such as geological surveying, radar detection/evasion, and aircraft design. These often involve high-frequency waves, which demand high-order methods to mitigate the dispersion error. We propose a high-order method for computing solutions to the variable-coefficient inhomogeneous Helmholtz equation in two dimensions on domains bounded by piecewise smooth curves of arbitrary shape with a finite number of boundary singularities at known locations.

We utilize compact finite difference (FD) schemes on regular structured grids to achieve high-order accuracy due to their efficiency and simplicity, as well as the capability to approximate variable-coefficient differential operators. In this work, a 4th-order compact FD scheme for the variable-coefficient Helmholtz equation on a Cartesian grid in 2D is derived and tested. The well known limitation of finite differences is that they lose accuracy when the boundary curve does not coincide with the discretization grid, which is a severe restriction on the geometry of the computational domain. Therefore, the algorithm presented in this work combines high-order FD schemes with the method of difference potentials (DP), which retains the efficiency of FD while allowing for boundary shapes that are not aligned with the grid without sacrificing the accuracy of the FD scheme.

Additionally, the theory of DP allows for the universal treatment of the boundary conditions. One of the significant contributions of this work is the development of an implementation that accommodates general boundary conditions (BCs). In particular, Robin BCs with discontinuous coefficients are studied, for which we introduce a piecewise parameterization of the boundary curve. Problems with discontinuities in the boundary data itself are also studied.

We observe that the design convergence rate suffers whenever the solution loses regularity due to the boundary conditions. This is because the FD scheme is only consistent for classical solutions of the PDE. For this reason, we implement the method of singularity subtraction as a means for restoring the design accuracy of the scheme in the presence of singularities at the boundary. While this method is well studied for low order methods and for problems in which singularities arise from the geometry (e.g., corners), we adapt it to our high-order scheme for curved boundaries via a conformal mapping and show that it can also be used to restore accuracy when the singularity arises from the BCs rather than the geometry.

Altogether, the proposed methodology for 2D boundary value problems is computationally efficient, easily handles a wide class of boundary conditions and boundary shapes that are not aligned with the discretization grid, and requires little modification for solving new problems.

© Copyright 2015 by Darrell Steven Britt, Jr.

All Rights Reserved

High-Order Accurate Solutions to the Helmholtz Equation
in the Presence of Boundary Singularities

by
Darrell Steven Britt, Jr.

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2015

APPROVED BY:

Alina Chertock

Moody Chu

Pierre Gremaud

Semyon Tsynkov
Chair of Advisory Committee

DEDICATION

This work is dedicated to my late father, Darrell Steven Britt, Sr. (August 27, 1964–May 4, 2010). He always reminded me that I could accomplish anything that I set my mind to. He was proud of me before I had accomplished a single thing in life. It is my honor to bear his name.

BIOGRAPHY

Darrell Steven Britt, Jr., was born to Darrell Steven Britt, Sr., and Dee Ann Britt in Edenton, North Carolina, on August 3, 1986. To avoid confusion, he goes by "Steven." He was raised primarily in the towns of Warner Robins and Perry, Georgia, during his formative years. To his own shame and despite his evident aptitude, he had no academic ambition or passion whatsoever until his senior year at Westfield High School in the 2003-2004 academic year. It was then that he discovered the beauty of Calculus, having read the entire textbook before the year was halfway finished – an act of previously unfamiliar and genuine scholastic interest. It was then that he decided to pursue a lifelong study of mathematics. He had no idea what he was getting himself into, but it was a wonderful decision.

Steven studied for one year at Macon State College before attending the Georgia Institute of Technology, where he earned a Bachelor's of Science in Applied Mathematics on May 5, 2007. Two weeks later, he was married to Danielle Britt. In August 2008, Steven enrolled in the graduate mathematics program at North Carolina State University in pursuit of a PhD in Applied Mathematics. In January 2009, he became a Research Assistant under the advisement of Dr. Semyon Tsynkov, who guided him in developing the enclosed work of this dissertation.

ACKNOWLEDGEMENTS

First and foremost, I give thanks to my Creator: God the Father and His Son Jesus Christ, who endowed man with the Intellect for such scholarly pursuits from the Beginning, who have personally sustained and guided me throughout this work, and to whom I offer every part of myself for their glory and pleasure, being as one made alive from the dead. May I grow in faith, obedience, and understanding of my God!

My heartfelt thanks goes to my loving and kind wife, Danielle, for giving me a compelling reason to be the best sort of man that I am capable of being in every respect – you deserve all of the effort that I have expended on your behalf, and more. May I honor and serve you in all I do, because you are the "bone of my bones and flesh of my flesh," and my most treasured possession.

I am indebted to my advisor, Dr. Semyon Tsynkov: a brilliant and patient man who has labored to help me reach my potential. At the early stage of my career at which we met, I could not have known what a blessing it would be to work with you. May I repay you double for all you have done for me.

Thanks to my family and friends who have made the duration of my graduate study the most enjoyable years of my life to date. I am fortunate to say that I need not list you all by name, because you know who you are by the love we have for one another. May the years to come be even better as we continue together in love and friendship!

Thank you to Dr. Eli Turkel, Dr. Sergey Petropavlovsky, and Dr. Michael Medvinsky for your intellectual contributions to this work and the spirit of collaboration that we share. May our success together endure!

Finally, I express my gratitude to the following funding sources, which have made my graduate study both possible and comfortable:

- US-Israel Binational Sciences Foundation grant 2008094
- US National Science Foundation grants DMS-0509695, DMS-0810963
- US Army Research Office grant W911NF-11-1-0384

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	ix
Chapter 1 Introduction	1
Chapter 2 Finite Difference Methods	5
2.1 Introduction	5
2.2 The Helmholtz equation in polar coordinates	6
2.2.1 One-dimensional example	7
2.2.2 A scheme for the two-dimensional polar Helmholtz equation	9
2.3 The variable-coefficient Helmholtz equation in 2D	12
2.4 Boundary conditions for a square domain	17
2.4.1 Dirichlet boundary conditions	17
2.4.2 Local Sommerfeld-type boundary conditions	19
2.5 Numerical verification of the Cartesian scheme	22
2.5.1 Dirichlet boundary condition	23
2.5.2 Local Sommerfeld-type boundary conditions	26
2.6 Auxiliary problems for the method of difference potentials	27
Chapter 3 Difference Potentials	30
3.1 Difference potentials and projections	31
3.1.1 The finite difference scheme and auxiliary problem	31
3.1.2 Grid sets and operators	32
3.2 Treatment of the boundary conditions	35
3.2.1 Equation-based extension to the grid boundary	35
3.2.2 Series representation of boundary functions	39
3.2.3 Applying the extension to form a linear system	40
3.2.4 The Robin boundary condition with smooth variable coefficients	43
3.2.5 The Robin boundary problem with discontinuous variable coefficients	46
3.2.6 Derivatives of the Chebyshev polynomials near the endpoints	52
3.2.7 Structure of the algorithm	55
3.3 Numerical results	56
3.3.1 Solution of multiple problems at low cost	56
3.3.2 Parameters of the computational setting	57
3.3.3 Smooth periodic Robin boundary condition via Fourier basis	61
3.3.4 Mixed Dirichlet/Neumann boundary conditions using Chebyshev basis	61
3.3.5 Piecewise smooth Robin boundary condition via Chebyshev basis	63
3.3.6 Reducing the number of basis functions for coarser grids	64
3.3.7 Increased efficiency by using different numbers of basis functions on individual segments of a split boundary	66
3.3.8 Inhomogeneous Problems	68
3.3.9 Loss of accuracy due to non-treatment of Chebyshev endpoints	69

3.3.10	Boundary data that lead to solutions with singularities	70
3.3.11	Effect of boundary data with varying degrees of continuity on convergence	74
3.4	Discussion and Remarks	76
Chapter 4	Singularity Subtraction	78
4.1	Constructing Singular Functions at the Boundary	79
4.1.1	Straight boundaries	80
4.1.2	Curved boundaries	86
4.1.3	Test problems	89
4.2	Solution by difference potentials in the presence of singularities	95
4.2.1	Regularized problem	95
4.2.2	Solution of multiple problems at low cost	96
4.3	Numerical simulations	97
4.3.1	Parameters of the computational setting	97
4.3.2	Test 1: Discontinuous Dirichlet boundary data	100
4.3.3	Test 2: Continuous Dirichlet boundary data with first derivative discontinuity.	102
4.3.4	Test 3: Continuous Dirichlet boundary conditions with second derivative discontinuity.	102
4.3.5	Test 4: Discontinuous Neumann boundary data	103
4.3.6	Test 5: Continuous Neumann boundary data with first derivative discontinuity.	104
4.3.7	Test 6: Continuous Neumann boundary data with second derivative discontinuity.	104
4.4	Discussion	105
Chapter 5	The Pollution Effect and Efficiency of the Algorithm	106
5.1	Numerical investigation of the Pollution Effect	107
5.1.1	Effect of the wavenumber on the pollution coefficient	109
5.1.2	Remarks and discussion	109
5.2	Choosing the Number of Basis Functions	110
5.2.1	Strategies	112
5.2.2	Numerical investigation of the adequate truncation tolerance	116
5.3	A Numerical Example Using the <i>a priori</i> Error Estimate	116
Chapter 6	Considerations for Future Work	120
6.1	Extensions of the 2D Algorithm	120
6.2	Extension of the algorithm to 3D	122
References	124
Appendices	132
Appendix A	Chebyshev coefficients for the product of two functions	133
Appendix B	Singularity Subtraction Test Problems 2 through 6	135

LIST OF TABLES

Table 2.1	Grid convergence and computation time data for example 1.	24
Table 2.2	Grid convergence and computation time data for example 2.	25
Table 2.3	Grid convergence and computation time data for example 3.	26
Table 2.4	Grid convergence and computation time data for example 4.	26
Table 2.5	Grid convergence and computation time data for example 5.	27
Table 2.6	Grid convergence and computation time data for example 6.	27
Table 3.1	Grid convergence and execution times for the smooth periodic Robin boundary condition (3.24). $N = 32$ in formula (3.17), and $\psi_n^{(0)}(\theta) = \psi_n^{(1)}(\theta) = e^{in\theta}$ in formula (3.16).	61
Table 3.2	Time to build the matrix \mathbf{Q} of (3.22).	62
Table 3.3	Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{ikx}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N = 67$	62
Table 3.4	Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{iky}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N = 67$	63
Table 3.5	Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) with $\alpha^{(1)}(\theta) = e^{\cos\theta}$, $\beta^{(1)}(\theta) = \arctan\theta + 1$ and $\alpha^{(2)}(\theta) = e^{2\sin\theta}$, $\beta^{(2)}(\theta) = 1$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16). The dimension of the Chebyshev basis on each arc is $N = 67$. 64	64
Table 3.6	Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) with $\alpha^{(1)}(\theta) = e^{\sin\theta}$, $\beta^{(1)}(\theta) = (\theta + 3)^2$ and $\alpha^{(2)}(\theta) = \log(\theta + 3)$, $\beta^{(2)}(\theta) = \sqrt{\theta + 3}$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16). The dimension of the Chebyshev basis on each arc is $N = 67$	64
Table 3.7	Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) of Table 3.5 with all the same parameters but fewer basis functions on coarser grids.	65
Table 3.8	Convergence data for the piecewise smooth Robin boundary condition (3.31) of Tables 3.5 and 3.7 with a split basis. The total number of basis functions used is $2(N_1 + N_2)$	67
Table 3.9	Inhomogeneous problem with right-hand side generated by the test solution $u = r^2 e^{ikx}$ with Dirichlet boundary conditions.	69
Table 3.10	Inhomogeneous problem with right-hand side generated by the test solution $u = r^3 e^{ikx}$ with Dirichlet boundary conditions.	69
Table 3.11	Computation of the same case as that in Table 3.5, but with no special treatment of the Chebyshev endpoints. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N=67$	70
Table 3.12	Grid convergence for boundary data with amplitude mismatch and either Dirichlet, Neumann, or mixed boundary conditions.	72

Table 3.13	Grid convergence for boundary data with direction mismatch (plane wave traveling in the y direction) and either Dirichlet, Neumann, or mixed boundary conditions.	72
Table 3.14	Grid convergence for boundary data with direction mismatch (plane wave traveling at 1 radian with respect to the positive x direction) and either Dirichlet, Neumann, or mixed boundary conditions.	73
Table 3.15	Grid convergence for boundary data with phase mismatch (waves out of phase by 0.7 radians) and either Dirichlet, Neumann, or mixed boundary conditions.	73
Table 3.16	Grid convergence for boundary data with no mismatch and either Dirichlet, Neumann, or mixed boundary conditions.	73
Table 3.17	Error and convergence rate for Dirichlet boundary conditions with 0^{th} up to 3^{rd} order discontinuity using boundary data (3.43-3.44)	75
Table 3.18	Error and convergence rate for Neumann boundary conditions with 0^{th} up to 3^{rd} order discontinuity using boundary data (3.43-3.44)	75
Table 4.1	Results for Dirichlet boundary data with discontinuity.	100
Table 4.2	Results for Dirichlet boundary data with discontinuity. The error is computed by comparison to the approximation (4.44) of the exact solution.	101
Table 4.3	Results for continuous Dirichlet boundary data with first derivative discontinuity.	102
Table 4.4	Results for continuous Dirichlet boundary data with second derivative discontinuity.	103
Table 4.5	Results for Neumann boundary data with discontinuity.	103
Table 4.6	Results for continuous Neumann boundary data with first derivative discontinuity.	104
Table 4.7	Results for continuous Neumann boundary data with second derivative discontinuity.	105
Table 5.1	Estimations of the pollution effect coefficient C_P for various simulations. . . .	108
Table 5.2	Estimation of the pollution coefficient C_P for $k = 3, \dots, 26$ on a 512×512 grid. . . .	110
Table 5.3	As the tolerances σ_1, σ_2 approach the error of the FD scheme, accuracy is lost. . . .	117
Table 5.4	The number of basis functions for the grids 64×64 and 128×128 is determined by the inequality (5.6), while for the finer grids it is determined by the truncation tolerance σ obtained from the average pollution coefficient from the previous computations of Table 5.2.	118

LIST OF FIGURES

Figure 2.1	Stencils of the compact scheme (2.30-2.32).	17
Figure 3.1	Interior and exterior grid subsets and the grid boundary.	32
Figure 4.1	Conformal mapping (4.19).	87
Figure 4.2	Function (4.37) on the disk $R = 1$. The parameters are $r_1 = 0.1$, $r_2 = 0.9$, $\varphi_1 = \pi/12$, $\varphi_2 = \pi/3$	93
Figure 5.1	Estimated values of the pollution coefficient C_P for $k = 3, \dots, 26$ from the 512×512 grid.	111
Figure 5.2	Real part of the numerically computed test solution $u = e^{ikx} + 2e^{i(2k)y}$ with wavenumber $k = 13.32$ on a 256×256 grid.	119

Chapter 1

Introduction

In the present work, we develop a high-order numerical method for solving linear elliptic PDEs with well-behaved variable coefficients on domains with piecewise-smooth boundaries and piecewise-smooth boundary conditions. We expect solutions to be predominantly classical, but some singularities may arise at known locations along the boundary due to either the geometry of the boundary shape (e.g., corners or cusps) or the boundary conditions (e.g., piecewise smooth with a known jump discontinuity). Our model equation for this pursuit will be the inhomogeneous Helmholtz equation,

$$\Delta u + k^2 u = f, \quad \mathbf{x} \in \Omega, \quad (1.1)$$

in which we permit the Laplacian operator to be replaced by a more general variable-coefficient operator and we allow for variation in the wavenumber, k . Variable coefficients may arise from the use of non-Cartesian geometries and may also occur naturally in problems of wave propagation through inhomogeneous media. In the case of inhomogeneous media, the equation (1.1) may more generally be written as follows:

$$\frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(b(x, y) \frac{\partial u}{\partial y} \right) + k^2(x, y) u = f, \quad (1.2)$$

where a, b , and k are smooth functions of (x, y) except perhaps along a known interface. The regions of smooth variation of these parameters represent materials with smoothly varying properties - for example, the density of a material may smoothly change when subjected to a heat source at one end. On the other hand, a jump condition at an interface would model contact between two different materials, such as an airplane which is composed of several different plastics and metals. These criteria are essential for simulations involving most real-world objects.

High-order accuracy is desirable in the numerical solution of wave propagation problems. It is known for the Helmholtz equation that in order to maintain a given level of error as the wavenumber k increases one must refine the computational grid faster than the wavelength $\lambda = 2\pi/k$ decreases. More precisely, the quantity $k^{p+1}h^p$ must remain constant, where h is the discretization size and p is the order of accuracy of the scheme. This is known as the dispersion error or pollution effect [1,2,3]. In order to avoid this undesirable behavior, the number of points per wavelength, $\sim 1/hk$, must grow proportional to $k^{1/p}$. It is evident then that employing a higher order of accuracy, p , slows this growth, meaning that higher-order methods will be considerably more efficient than lower-order methods for solving equation (1.1), especially on domains which are large relative to the wavelength. There are many applications in which the size of the object being studied is large in relation to the waves propagating through it, such as medical imaging (i.e., the propagation of ultrasound waves through the human body) and aircraft design (i.e., the propagation of electronic communication signals through the craft).

Finite difference (FD) schemes on regular-structured discretization grids provide the simplest and least expensive avenue for achieving high-order accuracy and can also easily be built for variable-coefficient equations; however, they are limited to classical solutions of the PDE and lose accuracy when the boundary shape is not aligned with the grid (e.g., a circular boundary on a Cartesian grid). The loss of accuracy in the presence of non-conforming boundaries is attributable to staircasing [4,5], and represents a substantial limitation. The primary alternatives to finite differences are finite element methods (FEM) and boundary element methods (BEM). Both of these methods easily accommodate difficult geometries, but each have limitations of their own. In particular, it is possible to build high-order approximations for arbitrary boundary shapes with FEM; however, achieving high order requires substantial modifications to the algorithm and also greatly increases the computational cost by requiring additional variables per grid node. In BEM, any change in the boundary conditions requires analysis of the equivalent boundary sources to ensure that the resulting integral equation is well-posed. Additionally, BEMs are limited to constant-coefficient PDEs since they rely on explicit knowledge of the fundamental solution. We propose a methodology which allows accurate computation by FD on grids which are not aligned with the boundary shape, and, in doing so, we do not compromise the utility or efficiency of FD in computing with high-order accuracy and for variable-coefficient equations. Furthermore, we develop a method for restoring accuracy in the case of near-boundary singularities.

In the following work, we employ the method of difference potentials (DP) developed by Ryaben'kii. The theory of difference potentials is related to the theory of Calderon operators, see [6,7]. A comprehensive account of the method of difference potentials can be found in the monograph [8] (see also [9]), and a brief account can also be found in [10,11]. Some of its other recent developments are presented in [12,13,14,15,16,17]. DP permits a broad range of

computational capabilities, see [17, 18], allowing for differential equations with variable coefficients and for both homogeneous and inhomogeneous equations. Additionally, in either the case of inhomogeneities or of variable properties of the medium, there is no requirement of having a gap between the region of inhomogeneity/variation and the boundary of the domain. The method also works for curvilinear boundaries and interfaces of arbitrary shape. In this work, we demonstrate its universal approach to the treatment of the boundary conditions. In doing so, we introduce a piecewise parameterization of the boundary curve in order to permit piecewise-smooth boundary conditions (i.e., those with discontinuities). While one natural use of the piecewise parameterization would be to solve problems with boundary shapes that require such a parameterization due to the geometry (such as domains with corners), the focus of the work at hand is on splitting the parameterization of a smooth boundary curve into multiple segments at points of discontinuity in the Robin coefficients or points of singularity in the boundary data themselves in order to restore the convergence order of the scheme. The method is designed so that it does not require the approximation of the boundary conditions on the grid and so is applicable to non-conforming boundaries. Moreover, boundary conditions of any type are allowed — not only the simplest Neumann or Dirichlet, but also, for example, mixed or Robin, even with variable and discontinuous coefficients. In doing so, the core of the numerical algorithm always remains the same. Changes in the boundary conditions are accommodated by making only minor modifications to the computational procedure (unlike methods based on boundary integral equations).

The class of problems we seek to solve are expected to have predominantly classical solutions; however, the flexibility of our approach to BCs can easily result in the solution becoming singular at the boundary, leading to a loss of consistency of the FD scheme and subsequently a loss of accuracy. In general, the solution may have lower regularity due to either a discontinuity in the coefficients of the boundary condition, a discontinuity in the boundary data themselves, or the geometry of the boundary curve (e.g., a cusp or corner), and our methodology can easily be applied to problems of each of these types. In such cases, a near-boundary singularity develops in the solution, which translates into a greatly diminished convergence rate from the FD scheme since the solution is no longer classical. The design convergence rate can be restored by subtracting several leading terms in the expansion of the solution near the singular point and then solving only for the remaining regular part. In developing this expansion, we follow an approach previously proposed in [19] and modify it so as to take into account that the segments of the boundary that meet at a given singular point may be curves rather than only straight lines, and do so in such a way that high-order accuracy is maintained. We also note that an earlier work that addresses singularities of the solution using difference potentials is presented in [20, 21, 22] but does not involve computations with high-order accuracy.

In Chapter 2, a fourth-order compact finite difference scheme for the variable-coefficient

Helmholtz equation (1.2) on a Cartesian grid is derived using the equation-based method [23,24], and this will serve as the solver for our subsequent work on difference potentials. In Chapter 3, we present a formulation of the method of difference potentials and detail a universal approach to treating boundary conditions [25]. We numerically demonstrate the limitations of the method by showing that the boundary conditions may give rise to singularities in the solution which cause a loss of accuracy. In Chapter 4, the method of [19] for treating near-boundary singularities is extended for use with high-order methods and implemented with the method of difference potentials to solve singular problems without degradation of the design convergence rate [26]. Chapter 5 discusses error estimation as it relates to the efficiency of the overall scheme, taking a second look at the numerical examples of the previous chapters and adding several numerical examples to give insight and guidance. Finally, Chapter 6 proposes future extensions of this work and discusses some of the challenges that will need to be addressed in doing so.

Chapter 2

Finite Difference Methods

2.1 Introduction

In this chapter, we build a fourth-order compact finite difference scheme for the variable-coefficient Helmholtz equation in two dimensions via the equation-based method [23, 24]. Traditional finite difference schemes achieve high order by extending the FD stencil outward in the coordinate directions, but this is undesirable because it increases the order of the difference equation beyond that of the underlying differential equation. In particular, this means that the difference equation will require more boundary conditions than those needed for the differential equation, meaning one must compensate by implementing purely numerical boundary conditions. One alternative is to use Padé expansions to replace the weight of the center node using a Bessel function [27], and this has been done for problems with Dirichlet and Neuman BCs [27, 28]; however, Pade approximations are limited to constant-coefficient equations. A more useful approach is to employ an equation-based method, in which the Helmholtz equation and its derivatives are used to find alternate representations of the truncation error terms of a classical second-order scheme, and these representations can then be sufficiently approximated on a compact 9-point stencil to increase the order of the scheme. It is important to note that, according to the well-known Lax theorem, the scheme needs only to be consistent on solutions to the equation [29]; therefore, the seeming limitation inherent in using the equation itself to build our scheme in fact represents no loss of generality. The 9-point (3×3) stencil is no wider in any coordinate direction than that of the underlying second-order scheme, eliminating the need for additional boundary constraints [27, 30, 31, 23, 24, 32] and also resulting in a narrower bandwidth of the FD matrix. Moreover, the equation-based approach can be used to construct high-order physical as well as radiation boundary conditions, as has been shown in [33]. While this technique will be used for the boundary conditions in this chapter, the choice of the BCs is not guided by physical boundary constraints but rather they are chosen to cater to the needs

of the auxiliary problems of the method of difference potentials used in subsequent chapters.

In Section 2.2, we demonstrate the equation-based method in simplified settings, beginning with the Helmholtz equation in 1D followed by a demonstration of how the method can also be used for the 2D Helmholtz equation in polar coordinates. As we progress in complexity, we next examine the general variable-coefficient equation (1.2) in Cartesian coordinates for 2D in Section 2.3 using a regular-structured Cartesian grid. Due to the requirement that finite difference schemes be aligned with the boundary shape of the domain, we subsequently apply the numerical method to interior problems on a square. On such a domain, the Helmholtz equation is known to experience resonances when the wavenumber k is such that $-k^2$ coincides with an eigenvalue of the Laplacian operator. This can be avoided by implementing Sommerfeld-type boundary conditions along the left and right edges of the square, with a simple Dirichlet boundary condition along the top and bottom edges. The high-order implementation of these boundary conditions is the subject of Section 2.4. Numerical simulations are performed in Section 2.5 to demonstrate the convergence rate. Finally, a discussion is given in Section 2.6 regarding the relationship of the finite difference schemes developed in this chapter to the method of difference potentials presented in Chapter 3, justifying the choice of boundary conditions for the finite difference problems in this chapter as well as the emphasis on the Cartesian scheme in particular.

2.2 The Helmholtz equation in polar coordinates

The Helmholtz equation in polar coordinates is given by

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + k^2 u = f. \quad (2.1)$$

The radial term of (2.1) is a particular example of the type of variable-coefficient differential operator that we treat in Section 2.3. As a demonstration of the equation-based method, we will first consider the Helmholtz equation in 1D, which is posed as the following inhomogeneous ODE:

$$\frac{1}{r} \frac{d}{dr} \left(r \frac{du}{dr} \right) + k^2 u = \frac{d^2 u}{dr^2} + \frac{1}{r} \frac{du}{dr} + k^2 u = f \quad (2.2)$$

where $f = f(r)$ is assumed given.

2.2.1 One-dimensional example

For convenience, we rewrite (2.2) by solving for the variable-coefficient term:

$$\frac{1}{r} \frac{d}{dr} \left(r \frac{du}{dr} \right) = f - k^2 u = F, \quad (2.3)$$

where $F := f - k^2 u$ is a formal right-hand side. We first approximate equation (2.3) with second order accuracy:

$$\frac{1}{r_m} \frac{1}{h} \left(r_{m+1/2} \frac{u_{m+1} - u_m}{h} - r_{m-1/2} \frac{u_m - u_{m-1}}{h} \right) = F_m \quad (2.4)$$

Analysis of the truncation error for scheme (2.4) shows that

$$\begin{aligned} \frac{1}{r_m} \frac{1}{h} \left(r_{m+1/2} \frac{u_{m+1} - u_m}{h} - r_{m-1/2} \frac{u_m - u_{m-1}}{h} \right) \\ = \frac{1}{r} \frac{d}{dr} \left(r \frac{du}{dr} \right) \Big|_m + \frac{h^2}{12} \left(u_m^{(4)} + \frac{2}{r} u_m^{(3)} \right) + \mathcal{O}(h^4) \end{aligned} \quad (2.5)$$

Consequently, to achieve fourth order accuracy, we need to eliminate the $\mathcal{O}(h^2)$ term that contains $u^{(4)}$ and $u^{(3)}$ on the right-hand side of (2.5). We begin by rearranging the ODE (2.3) to isolate the second derivative term:

$$\frac{d^2 u}{dr^2} = F - \frac{1}{r} \frac{du}{dr}. \quad (2.6)$$

We then differentiate this expression twice to obtain

$$\frac{d^3 u}{dr^3} = F' - \frac{1}{r} \frac{d^2 u}{dr^2} + \frac{1}{r^2} \frac{du}{dr} \quad (2.7)$$

and

$$\frac{d^4 u}{dr^4} = F'' - \frac{1}{r} \frac{d^3 u}{dr^3} + \frac{2}{r^2} \frac{d^2 u}{dr^2} - \frac{2}{r^3} \frac{du}{dr}. \quad (2.8)$$

Furthermore, we substitute equation (2.6) into (2.7) and simplify as follows:

$$\begin{aligned} \frac{d^3 u}{dr^3} &= F' - \frac{1}{r} \frac{d^2 u}{dr^2} + \frac{1}{r^2} \frac{du}{dr} \\ &= F' - \frac{1}{r} \left(F - \frac{1}{r} \frac{du}{dr} \right) + \frac{1}{r^2} \frac{du}{dr} \\ &= F' - \frac{1}{r} F + \frac{2}{r} \frac{du}{dr}. \end{aligned} \quad (2.9)$$

Likewise, we may now substitute expressions (2.6) and (2.9) into (2.8), which ultimately yields

$$\frac{d^4 u}{dr^4} = F'' - \frac{1}{r} F' + \frac{3}{r^2} F - \frac{4}{r^3} \frac{du}{dr}. \quad (2.10)$$

We have now achieved alternate representations of the terms $u^{(3)}$ and $u^{(4)}$ which appear in the $\mathcal{O}(h^2)$ term of (2.5), so that we can now replace the term $u^{(4)} + \frac{2}{r} u^{(3)}$ as follows using (2.9-2.10):

$$\begin{aligned} \frac{d^4 u}{dr^4} + \frac{2}{r} \frac{d^3 u}{dr^3} &= F'' + \frac{1}{r} F' + \frac{1}{r^2} \frac{d^2 u}{dr^2} - \frac{1}{r^3} \frac{du}{dr} \\ &= F'' + \frac{1}{r} F' + \frac{1}{r^2} \left(\frac{d^2 u}{dr^2} - \frac{1}{r} \frac{du}{dr} \right) \\ &= F'' + \frac{1}{r} F' + \frac{1}{r^2} \left(F - \frac{2}{r} \frac{du}{dr} \right), \end{aligned}$$

from which we obtain, via formula (2.5), a fourth-order approximation:

$$\begin{aligned} \frac{1}{r_m} \frac{1}{h} \left(r_{m+1/2} \frac{u_{m+1} - u_m}{h} - r_{m-1/2} \frac{u_m - u_{m-1}}{h} \right) \\ - \frac{h^2}{12} \left(F'' + \frac{1}{r} F' + \frac{1}{r^2} F - \frac{2}{r^3} \frac{du}{dr} \right) \Big|_m \\ = \frac{1}{r} \frac{d}{dr} \left(r \frac{du}{dr} \right) \Big|_m + \mathcal{O}(h^4). \end{aligned} \quad (2.11)$$

Relation (2.11) yields the following approximation for the original ODE (2.2), provided that k is constant:

$$\begin{aligned} \frac{1}{r_m} \frac{1}{h} \left(r_{m+1/2} \frac{u_{m+1} - u_m}{h} - r_{m-1/2} \frac{u_m - u_{m-1}}{h} \right) + k^2 u_m \\ - \frac{h^2}{12} \left(F'' + \frac{1}{r} F' + \frac{1}{r^2} F - \frac{2}{r^3} \frac{du}{dr} \right) \Big|_m = f_m, \end{aligned} \quad (2.12)$$

where the equation (2.2) was used to replace the truncation error terms of scheme (2.4) inside the term $\sim \frac{h^2}{12}$. Equation (2.12) is not quite a true finite-difference scheme because the terms F, F' , and F'' contain continuous derivatives of f and of u . To transform (2.12) into a fourth-order finite difference scheme, we first realize that the term multiplied by $\frac{h^2}{12}$ on the left-hand side of (2.12) does not need to be evaluated exactly: it is sufficient to have it approximated with second order accuracy since it is multiplied by the factor of $\frac{h^2}{12}$. Clearly this can be done for the terms f, f', f'', u, u' and u'' by means of the standard central differences on a 3-node stencil. If $f(r)$ is known analytically, then one may use the exact derivatives of f . In either case, once this is done the approximating relation (2.12) becomes a fourth-order accurate finite difference scheme for equation (2.2) while still maintaining a compact 3-node stencil.

As we will see in the next section, the equation-based method can also be used to build high-order schemes in 2D for the polar Helmholtz equation on a compact 3×3 stencil. Though the expressions become more complicated, the key idea of using the derivatives of the equation to obtain alternate representations of the terms of the truncation error which do not require a larger stencil remains the same.

2.2.2 A scheme for the two-dimensional polar Helmholtz equation

The key consideration that enables us to extend the methodology of Section 2.2.1 to the two-dimensional equation (2.1) is that the fourth order accurate approximations will be built independently for the individual second order differential operators of the Laplacian. Hence, we write the following formal ODEs based on equation (2.1):

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) = F_r \equiv f - k^2 u - \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \quad (2.13a)$$

$$\frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} = F_\theta \equiv f - k^2 u - \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) \quad (2.13b)$$

Equations (2.13a) and (2.13b) are each identical to (2.3) up to the notation. Consequently, we may continue to use formula (2.11) to obtain a fourth order accurate approximation of the radial part of the Laplacian in (2.1):

$$\begin{aligned} \frac{1}{r_m} \frac{1}{h_r} \left(r_{m+1/2} \frac{u_{m+1,n} - u_{m,n}}{h_r} - r_{m-1/2} \frac{u_{m,n} - u_{m-1,n}}{h_r} \right) \\ - \frac{h_r^2}{12} \left(\frac{\partial^2 F_r}{\partial r^2} + \frac{1}{r} \frac{\partial F_r}{\partial r} + \frac{1}{r^2} F_r - \frac{2}{r^3} \frac{\partial u}{\partial r} \right) \Big|_{m,n} \\ = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) \Big|_{m,n} + \mathcal{O}(h_r^4). \end{aligned} \quad (2.14)$$

Relation (2.14) is different from its “parent” relation (2.11) in that the auxiliary right-hand side F_r also contains the second derivative with respect to θ , see formula (2.13a) so that

$$\begin{aligned} \frac{\partial F_r}{\partial r} &= \frac{\partial f}{\partial r} - k^2 \frac{\partial u}{\partial r} - \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right), \\ \frac{\partial^2 F_r}{\partial r^2} &= \frac{\partial^2 f}{\partial r^2} - k^2 \frac{\partial^2 u}{\partial r^2} - \frac{\partial^2}{\partial r^2} \left(\frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right). \end{aligned} \quad (2.15)$$

Next, approximating the derivatives of u within (2.15) with second order accuracy via central differences, we have

$$\begin{aligned}
F_r|_{m,n} &= f_{m,n} - k^2 u_{m,n} - \frac{1}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} + \mathcal{O}(h_\theta^2), \\
\frac{\partial F_r}{\partial r}|_{m,n} &= \frac{\partial f}{\partial r}|_{m,n} - k^2 \frac{u_{m+1,n} - u_{m-1,n}}{2h_r} \\
&\quad - \frac{1}{2h_r} \left(\frac{1}{r_{m+1}^2} \frac{u_{m+1,n+1} - 2u_{m+1,n} + u_{m+1,n-1}}{h_\theta^2} \right. \\
&\quad \left. - \frac{1}{r_{m-1}^2} \frac{u_{m-1,n+1} - 2u_{m-1,n} + u_{m-1,n-1}}{h_\theta^2} \right) + \mathcal{O}(h_r^2 + h_\theta^2), \\
\frac{\partial^2 F_r}{\partial r^2}|_{m,n} &= \frac{\partial^2 f}{\partial r^2}|_{m,n} - k^2 \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h_r^2} \\
&\quad - \frac{1}{h_r^2} \left(\frac{1}{r_{m+1}^2} \frac{u_{m+1,n+1} - 2u_{m+1,n} + u_{m+1,n-1}}{h_\theta^2} \right. \\
&\quad - \frac{2}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} \\
&\quad \left. + \frac{1}{r_{m-1}^2} \frac{u_{m-1,n+1} - 2u_{m-1,n} + u_{m-1,n-1}}{h_\theta^2} \right) + \mathcal{O}(h_r^2 + h_\theta^2).
\end{aligned} \tag{2.16}$$

Substituting expressions (2.16) into (2.14) and also using the central difference approximation

$$\frac{\partial u}{\partial r}|_{m,n} = \frac{u_{m+1,n} - u_{m-1,n}}{2h_r} + \mathcal{O}(h_r^2), \tag{2.17}$$

we obtain a fourth order accurate finite difference approximation of $\frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right)$ on a compact 3×3 stencil. Note again that the derivatives of f in formulae (2.16) can be computed either analytically or also numerically by central differences, depending on how the right-hand side is defined.

The treatment of the second derivative with respect to θ , which will be based on equation (2.13b), is even more straightforward. We begin with the standard second order accurate central difference scheme:

$$\frac{1}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} = F_\theta|_{m,n}$$

The analysis of its truncation error shows that

$$\frac{1}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} = \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{h_\theta^2}{12} \left(\frac{1}{r^2} \frac{\partial^4 u}{\partial \theta^4} \right) + \mathcal{O}(h_\theta^4).$$

Consequently,

$$\frac{1}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} - \frac{h_\theta^2}{12} \frac{\partial^2 F_\theta}{\partial \theta^2} \Big|_{m,n} = \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \Big|_{m,n} + \mathcal{O}(h_\theta^4). \quad (2.18)$$

From differentiating (2.13b), we have

$$\frac{\partial^2 F_\theta}{\partial \theta^2} = \frac{\partial^2 f}{\partial \theta^2} - k^2 \frac{\partial^2 u}{\partial \theta^2} - \frac{1}{r} \frac{\partial^3}{\partial r \partial \theta^2} \left(r \frac{\partial u}{\partial r} \right). \quad (2.19)$$

Using central differences to approximate the terms of (2.19) yields:

$$\begin{aligned} \frac{\partial^2 F_\theta}{\partial \theta^2} \Big|_{m,n} &= \frac{\partial^2 f}{\partial \theta^2} \Big|_{m,n} - k^2 \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} \\ &\quad - \frac{1}{r_m} \frac{1}{h_r^2} \left(r_{m+1/2} \frac{u_{m+1,n+1} - 2u_{m+1,n} + u_{m+1,n-1}}{h_\theta^2} \right. \\ &\quad \left. - 2r_m \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} \right. \\ &\quad \left. + r_{m-1/2} \frac{u_{m-1,n+1} - 2u_{m-1,n} + u_{m-1,n-1}}{h_\theta^2} \right) + \mathcal{O}(h_r^2 + h_\theta^2). \end{aligned} \quad (2.20)$$

Substituting expression (2.20) into (2.18), we obtain a fourth order accurate finite-difference approximation of $\frac{\partial^2 u}{\partial \theta^2}$ on a compact 3×3 stencil.

The overall fourth order accurate compact scheme for the Helmholtz equation (2.1) is then obtained by combining (2.14) and (2.18):

$$\begin{aligned} &\frac{1}{r_m} \frac{1}{h_r} \left(r_{m+1/2} \frac{u_{m+1,n} - u_{m,n}}{h_r} - r_{m-1/2} \frac{u_{m,n} - u_{m-1,n}}{h_r} \right) \\ &\quad + \frac{1}{r_m^2} \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h_\theta^2} + k^2 u_{m,n} \\ &\quad - \frac{h_r^2}{12} \left(\frac{\partial^2 F_r}{\partial r^2} + \frac{1}{r} \frac{\partial F_r}{\partial r} + \frac{1}{r^2} F_r - \frac{2}{r^3} \frac{\partial u}{\partial r} \right) \Big|_{m,n} - \frac{h_\theta^2}{12} \frac{\partial^2 F_\theta}{\partial \theta^2} \Big|_{m,n} = f_{m,n}, \end{aligned} \quad (2.21)$$

where the correction terms $\sim \frac{h_r^2}{12}$ and $\sim \frac{h_\theta^2}{12}$ are to be evaluated according to (2.16), (2.17), and (2.20).

This same procedure of eliminating the truncation error by the equation-based method can also be used to build high-order schemes for other equations. In particular, the goal of the next section is to do the same for a more general form of the Helmholtz equation in Cartesian coordinates.

2.3 The variable-coefficient Helmholtz equation in 2D

We now construct a compact fourth order accurate finite difference scheme for the variable coefficient Helmholtz equation (1.2). As in Section 2.2.2, we begin by reformulating equation (1.2) as formal ODEs in order to treat the second-order differential terms of the Laplacian separately:

$$\frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right) = F_x \stackrel{\text{def}}{=} f - k^2 u - \frac{\partial}{\partial y} \left(b(x, y) \frac{\partial u}{\partial y} \right) \quad (2.22a)$$

and

$$\frac{\partial}{\partial y} \left(b(x, y) \frac{\partial u}{\partial y} \right) = F_y \stackrel{\text{def}}{=} f - k^2 u - \frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right). \quad (2.22b)$$

Using the same step size in each direction, $h = h_x = h_y$, we next approximate the left-hand sides of (2.22a) and (2.22b) at the grid node (m, n) with second order accuracy as follows:

$$\frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right) = \frac{1}{h} \left(a_{m+\frac{1}{2},n} \frac{u_{m+1,n} - u_{m,n}}{h} - a_{m-\frac{1}{2},n} \frac{u_{m,n} - u_{m-1,n}}{h} \right) + \mathcal{O}(h^2) \quad (2.23a)$$

and

$$\frac{\partial}{\partial y} \left(b(x, y) \frac{\partial u}{\partial y} \right) = \frac{1}{h} \left(b_{m,n+\frac{1}{2}} \frac{u_{m,n+1} - u_{m,n}}{h} - b_{m,n-\frac{1}{2}} \frac{u_{m,n} - u_{m,n-1}}{h} \right) + \mathcal{O}(h^2). \quad (2.23b)$$

Adding (2.23a) and (2.23b) and then approximating the non-differentiated term and the right-hand side of (1.2) as $(k^2 u)_{m,n}$ and $f_{m,n}$ respectively, we obtain a second order approximation of (1.2). Our aim is to extend (2.23) to a fourth order accurate approximation. Analysis of the truncation error for the finite differences of (2.23) yields the following:

$$\frac{1}{h} \left(a_{m+\frac{1}{2},n} \frac{u_{m+1,n} - u_{m,n}}{h} - a_{m-\frac{1}{2},n} \frac{u_{m,n} - u_{m-1,n}}{h} \right) \quad (2.24a)$$

$$= \frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right) + \frac{h^2}{12} \left(a u_{xxxx} + 2a_x u_{xxx} + \frac{3a_{xx} u_{xx}}{2} + \frac{a_{xxx} u_x}{2} \right) + \mathcal{O}(h^4),$$

$$\frac{1}{h} \left(b_{m,n+\frac{1}{2}} \frac{u_{m,n+1} - u_{m,n}}{h} - b_{m,n-\frac{1}{2}} \frac{u_{m,n} - u_{m,n-1}}{h} \right) \quad (2.24b)$$

$$= \frac{\partial}{\partial y} \left(b(x, y) \frac{\partial u}{\partial y} \right) + \frac{h^2}{12} \left(b u_{yyyy} + 2b_y u_{yyy} + \frac{3b_{yy} u_{yy}}{2} + \frac{b_{yyy} u_y}{2} \right) + \mathcal{O}(h^4).$$

In order to eliminate the $\mathcal{O}(h^2)$ error terms in (2.24a), we first differentiate the ODE (2.22a)

twice in x and obtain a system of three equations with respect to u_{xx} , u_{xxx} , and u_{xxxx} :

$$\begin{aligned} a_x u_x + a u_{xx} &= F_x, \\ a_{xx} u_x + 2a_x u_{xx} + a u_{xxx} &= \frac{\partial F_x}{\partial x}, \\ a_{xxx} u_x + 3a_{xx} u_{xx} + 3a_x u_{xxx} + a u_{xxxx} &= \frac{\partial^2 F_x}{\partial x^2}. \end{aligned} \quad (2.25)$$

Solving each equation of (2.25) for the highest derivative of u , we obtain

$$u_{xx} = \frac{1}{a}(F_x - a_x u_x), \quad (2.26a)$$

$$u_{xxx} = \frac{1}{a} \left(\frac{\partial F_x}{\partial x} - a_{xx} u_x - 2a_x u_{xx} \right), \quad (2.26b)$$

$$u_{xxxx} = \frac{1}{a} \left(\frac{\partial^2 F_x}{\partial x^2} - a_{xxx} u_x - 3a_{xx} u_{xx} - 3a_x u_{xxx} \right). \quad (2.26c)$$

Substituting (2.26a) into (2.26b) and also substituting (2.26a-2.26b) into (2.26c), we arrive at equations that contain only first derivatives of u on the right-hand side:

$$u_{xx} = \frac{1}{a}(F_x - a_x u_x), \quad (2.27a)$$

$$u_{xxx} = \frac{1}{a} \left(\frac{\partial F_x}{\partial x} - a_{xx} u_x - \frac{2a_x}{a}(F_x - a_x u_x) \right), \quad (2.27b)$$

$$\begin{aligned} u_{xxxx} &= \frac{1}{a} \left[\frac{\partial^2 F_x}{\partial x^2} - a_{xxx} u_x - \frac{3a_{xx}}{a}(F_x - a_x u_x) \right] \\ &\quad - \frac{3a_x}{a^2} \left[\frac{\partial F_x}{\partial x} - a_{xx} u_x - \frac{2a_x}{a}(F_x - a_x u_x) \right]. \end{aligned} \quad (2.27c)$$

By substituting expressions (2.26a), (2.27b), and (2.27c) for u_{xx} , u_{xxx} , and u_{xxxx} , respectively, into the $\mathcal{O}(h^2)$ terms on the right-hand side of (2.24a), we have

$$\begin{aligned} a u_{xxxx} + 2a_x u_{xxx} + \frac{3a_{xx} u_{xx}}{2} + \frac{a_{xxx} u_x}{2} &= \frac{\partial^2 F_x}{\partial x^2} - \frac{a_x}{a} \frac{\partial F_x}{\partial x} + \left(\frac{2a_x^2}{a^2} - \frac{3a_{xx}}{2a} \right) F_x \\ &\quad + \left(-\frac{a_{xxx}}{2} + \frac{5a_{xx} a_x}{2a} - \frac{2a_x^3}{a^2} \right) u_x. \end{aligned} \quad (2.28)$$

To achieve overall fourth order accuracy for (2.24a), it is sufficient to approximate the terms multiplied by $\frac{h^2}{12}$ (i.e., the right-hand side of (2.28)) with second order accuracy. For simplicity, assume that all derivatives of a and b on the right-hand side of (2.28) are known analytically.¹ Next, we differentiate formula (2.22a) to obtain formulae for $\frac{\partial F_x}{\partial x}$ and $\frac{\partial^2 F_x}{\partial x^2}$:

¹Otherwise, we can also replace them by finite differences if a and b can be sampled on the grid, although this may require a larger stencil.

$$\begin{aligned}
F_x &= f - k^2 u - b_y u_y - b u_{yy}, \\
\frac{\partial F_x}{\partial x} &= f_x - [b_{xy} u_y + b_y u_{xy} + b_x u_{yy} + b u_{yyx} + (k^2 u)_x], \\
\frac{\partial^2 F_x}{\partial x^2} &= f_{xx} - [b_{yxx} u_y + 2b_{yx} u_{yx} + b_y u_{yxx} + b_{xx} u_{yy} + 2b_x u_{yyx} + b u_{yyxx} + (k^2 u)_{xx}].
\end{aligned} \tag{2.29}$$

We use central differences on the standard five-node stencil to approximate $(k^2 u)_x$, $(k^2 u)_{xx}$, u_y , u_{yy} , f_x , and f_{xx} in (2.29) with second-order accuracy. We approximate the remaining terms u_{xy} , u_{yyx} , u_{xxy} , and u_{xxyy} on a compact 3×3 stencil (which contains the four additional corner nodes) also with second order accuracy as follows:

$$\begin{aligned}
u_{xy} &= \frac{1}{2h} \left(\frac{u_{m+1,n+1} - u_{m-1,n+1}}{2h} - \frac{u_{m+1,n-1} - u_{m-1,n-1}}{2h} \right) + \mathcal{O}(h^2), \\
u_{xxy} &= \frac{1}{h^2} \left(\frac{u_{m+1,n+1} - u_{m+1,n-1}}{2h} + \frac{u_{m-1,n+1} - u_{m-1,n-1}}{2h} - 2 \frac{u_{m,n+1} - u_{m,n-1}}{2h} \right) + \mathcal{O}(h^2), \\
u_{yyx} &= \frac{1}{h^2} \left(\frac{u_{m+1,n+1} - u_{m-1,n+1}}{2h} + \frac{u_{m+1,n-1} - u_{m-1,n-1}}{2h} - 2 \frac{u_{m+1,n} - u_{m-1,n}}{2h} \right) + \mathcal{O}(h^2), \\
u_{xxyy} &= \frac{1}{h^2} \left(\frac{u_{m+1,n+1} + u_{m+1,n-1} - 2u_{m+1,n}}{h^2} + \frac{u_{m-1,n+1} + u_{m-1,n-1} - 2u_{m-1,n}}{h^2} \right. \\
&\quad \left. - 2 \frac{u_{m,n+1} + u_{m,n-1} + 2u_{m,n}}{h^2} \right) + \mathcal{O}(h^2).
\end{aligned}$$

Altogether, we obtain a second order accurate approximation of all the terms on the right-hand sides of equalities (2.29) on a 3×3 stencil:

$$\begin{aligned}
F_x &= f_{m,n} - (k^2 u)_{m,n} - \frac{b_y}{2h} (u_{m,n+1} - u_{m,n-1}) \\
&\quad - b \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h^2} + \mathcal{O}(h^2),
\end{aligned} \tag{2.30a}$$

$$\begin{aligned}
\frac{\partial F_x}{\partial x} &= f_x - [b_{xy} u_y + b_y u_{xy} + b_x u_{yy} + b u_{yyx} + (k^2 u)_x] \\
&= \frac{f_{m+1,n} - f_{m-1,n}}{2h} - \left[\frac{b_{xy}}{2h} (u_{m,n+1} - u_{m,n-1}) \right. \\
&\quad + \frac{b_y}{4h^2} (u_{m+1,n+1} - u_{m-1,n+1} - u_{m+1,n-1} + u_{m-1,n-1}) \\
&\quad + \frac{b_x}{h^2} (u_{m,n+1} + u_{m,n-1} - 2u_{m,n}) \\
&\quad + \frac{b}{2h^3} (u_{m+1,n+1} - u_{m-1,n+1} + u_{m+1,n-1} - u_{m-1,n-1} \\
&\quad \left. - 2(u_{m+1,n} - u_{m-1,n})) + \frac{(k^2 u)_{m+1,n} - (k^2 u)_{m-1,n}}{2h} \right] + \mathcal{O}(h^2),
\end{aligned} \tag{2.30b}$$

$$\frac{\partial^2 F_x}{\partial x^2} = f_{xx} - [b_{yxx} u_y + 2b_{yx} u_{yx} + b_y u_{yxx} + b_{xx} u_{yy} \tag{2.30c}$$

$$\begin{aligned}
& + 2b_x u_{yyx} + b u_{yyxx} + (k^2 u)_{xx}] \\
= & \frac{f_{m+1,n} + f_{m-1,n} - 2f_{m,n}}{h^2} - \left[\frac{b_{yxx}}{2h} (u_{m,n+1} - u_{m,n-1}) \right. \\
& + \frac{b_{yx}}{2h^2} (u_{m+1,n+1} - u_{m-1,n+1} - u_{m+1,n-1} + u_{m-1,n-1}) \\
& + \frac{b_y}{2h^3} (u_{m+1,n+1} - u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} \\
& - 2(u_{m,n+1} - u_{m,n-1})) + \frac{b_{xx}}{h^2} (u_{m,n+1} + u_{m,n-1} - 2u_{m,n}) \\
& + \frac{b_x}{h^3} (u_{m+1,n+1} - u_{m-1,n+1} + u_{m+1,n-1} - u_{m-1,n-1} - 2(u_{m+1,n} - u_{m-1,n})) \\
& + \frac{b}{h^4} (u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} + u_{m-1,n-1} + 4u_{m,n} \\
& - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})) \\
& \left. + \frac{(k^2 u)_{m+1,n} - 2(k^2 u)_{m,n} + (k^2 u)_{m-1,n}}{h^2} \right] + \mathcal{O}(h^2).
\end{aligned}$$

As there is a complete symmetry between the derivatives in the x and y directions, the entire previous derivation can be easily duplicated in the y direction. Namely, we start with differentiating equation (2.22b) twice in y [cf. formula (2.25)], then we express the $\mathcal{O}(h^2)$ term on the right hand side of (2.24b) via u_y , F_y , $\frac{\partial F_y}{\partial y}$, and $\frac{\partial^2 F_y}{\partial y^2}$ [cf. formula (2.28)], and, in order to obtain a compact discretization, approximate the resulting terms with second order accuracy on a 3×3 stencil:

$$\begin{aligned}
F_y = & f_{m,n} - (k^2)u_{m,n} - \frac{a_x}{2h}(u_{m+1,n} - u_{m-1,n}) \\
& - a \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} + \mathcal{O}(h^2),
\end{aligned} \tag{2.31a}$$

$$\begin{aligned}
\frac{\partial F_y}{\partial y} = & f_y - [a_{xy}u_x + a_x u_{xy} + a_y u_{xx} + a u_{xxy} + (k^2 u)_y] \\
= & \frac{f_{m,n+1} - f_{m,n-1}}{2h} - \left[\frac{a_{xy}}{2h} (u_{m+1,n} - u_{m-1,n}) \right. \\
& + \frac{a_x}{4h^2} (u_{m+1,n+1} - u_{m-1,n+1} - u_{m+1,n-1} + u_{m-1,n-1}) \\
& + \frac{a_y}{h^2} (u_{m+1,n} + u_{m-1,n} - 2u_{m,n}) \\
& + \frac{a}{2h^3} (u_{m+1,n+1} - u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} \\
& - 2(u_{m,n+1} - u_{m,n-1})) + \frac{(k^2 u)_{m,n+1} - (k^2 u)_{m,n-1}}{2h} \left. \right] + \mathcal{O}(h^2),
\end{aligned} \tag{2.31b}$$

$$\begin{aligned}
\frac{\partial^2 F_y}{\partial y^2} = & f_{yy} - [a_{yyy}u_x + 2a_{xy}u_{xy} + a_x u_{xxy} + a_{yy}u_{xx} \\
& + 2a_y u_{xxy} + a u_{xxyy} + (k^2 u)_{yy}]
\end{aligned} \tag{2.31c}$$

$$\begin{aligned}
&= \frac{f_{m,n+1} + f_{m,n-1} - 2f_{m,n}}{h^2} - \left[\frac{a_{xyy}}{2h} (u_{m+1,n} - u_{m-1,n}) \right. \\
&\quad + \frac{a_{xy}}{2h^2} (u_{m+1,n+1} - u_{m-1,n+1} - u_{m+1,n-1} + u_{m-1,n-1}) \\
&\quad + \frac{a_x}{2h^3} (u_{m+1,n+1} - u_{m-1,n+1} + u_{m+1,n-1} - u_{m-1,n-1} \\
&\quad - 2(u_{m+1,n} - u_{m-1,n})) + \frac{a_{yy}}{h^2} (u_{m+1,n} + u_{m-1,n} - 2u_{m,n}) \\
&\quad + \frac{a_y}{h^3} (u_{m+1,n+1} - u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} - 2(u_{m,n+1} - u_{m,n-1})) \\
&\quad + \frac{b}{h^4} (u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} + u_{m-1,n-1} + 4u_{m,n} \\
&\quad - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})) \\
&\quad \left. + \frac{(k^2 u)_{m,n+1} - 2(k^2 u)_{m,n} + (k^2 u)_{m,n-1}}{h^2} \right] + \mathcal{O}(h^2).
\end{aligned}$$

Finally, assembling all the terms, we obtain a fourth order accurate approximation for equation (1.2) on a compact 3×3 stencil:

$$\begin{aligned}
&\frac{1}{h} \left(a_{m+\frac{1}{2},n} \frac{u_{m+1,n} - u_{m,n}}{h} - a_{m-\frac{1}{2},n} \frac{u_{m,n} - u_{m-1,n}}{h} \right) \\
&+ \frac{1}{h} \left(b_{m,n+\frac{1}{2}} \frac{u_{m,n+1} - u_{m,n}}{h} - b_{m,n-\frac{1}{2}} \frac{u_{m,n} - u_{m,n-1}}{h} \right) \\
&\quad - \frac{h^2}{12} \left(\frac{\partial^2 F_x}{\partial x^2} - \frac{a_x}{a} \frac{\partial F_x}{\partial x} + \left[\frac{2a_x^2}{a^2} - \frac{3a_{xx}}{2a} \right] F_x \right. \\
&\quad \left. + \left[-\frac{a_{xxx}}{2} + \frac{5a_{xx}a_x}{2a} - \frac{2a_x^3}{a^2} \right] \frac{u_{m+1,n} - u_{m-1,n}}{2h} \right) \Big|_{m,n} \\
&\quad - \frac{h^2}{12} \left(\frac{\partial^2 F_y}{\partial y^2} - \frac{b_y}{b} \frac{\partial F_y}{\partial y} + \left[\frac{2b_y^2}{b^2} - \frac{3b_{yy}}{2b} \right] F_y \right. \\
&\quad \left. + \left[-\frac{b_{yyy}}{2} + \frac{5b_{yy}b_y}{2b} - \frac{2b_y^3}{b^2} \right] \frac{u_{m,n+1} - u_{m,n-1}}{2h} \right) \Big|_{m,n} + (k^2 u)_{m,n} = f_{m,n}
\end{aligned} \tag{2.32}$$

In formula (2.32), the terms in parentheses premultiplied by $\frac{h^2}{12}$ are evaluated on the grid with second order accuracy using formulae (2.30) and (2.31). Scheme (2.30-2.32) is built on a square-cell Cartesian grid with step size h . It can be decomposed into a 9-node (3×3) compact stencil operating on u and a 5-node stencil operating on f , see Figure 2.1.

In Section 2.5, the scheme (2.30-2.32) is applied to several variable-coefficient problems; however, for Chapters 3 and 4, only constant-coefficient problems will be solved. In the case of

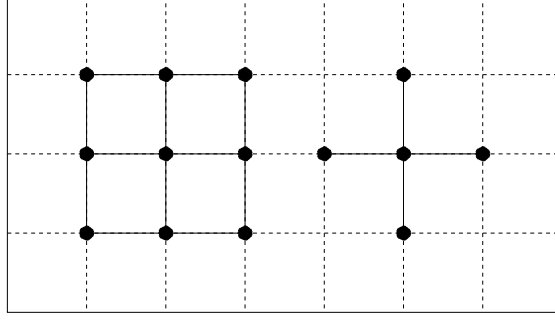


Figure 2.1: Stencils of the compact scheme (2.30-2.32).

constant coefficients, the scheme (2.30-2.32) assumes a much simpler form:

$$\begin{aligned}
& \frac{1}{h^2} (u_{m+1,n} + u_{m,n+1} + u_{m-1,n} + u_{m,n-1} - 4u_{m,n}) \\
& + \frac{1}{6h^2} [u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} + 4u_{m,n} \\
& - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})] \\
& + \frac{k^2}{12} (u_{m+1,n} + u_{m,n+1} + 8u_{m,n} + u_{m-1,n} + u_{m,n-1}) \\
& = f_{m,n} + \frac{1}{12} (f_{m+1,n} + f_{m,n+1} - 4f_{m,n} + f_{m-1,n} + f_{m,n-1}).
\end{aligned} \tag{2.33}$$

We note also that a sixth order accurate scheme is constructed for the constant-coefficient Helmholtz equation in [31] using the same 9-node compact stencil on the left-hand side, and a sixth order compact scheme is built for the Helmholtz equation with a variable wavenumber k in [32].

2.4 Boundary conditions for a square domain

In order to maintain high order accuracy, it is necessary that all boundary conditions be accurate to the same order as the interior scheme, see [34].

2.4.1 Dirichlet boundary conditions

We first consider the variable coefficient Helmholtz equation with constant wavenumber k on a square domain of side length s , $\Omega = \{(x, y) | -\frac{s}{2} < x < \frac{s}{2}, -\frac{s}{2} < y < \frac{s}{2}\}$:

$$-\frac{\partial}{\partial x} \left(a \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(b \frac{\partial u}{\partial y} \right) - k^2 u = -f, \quad (x, y) \in \Omega, \tag{2.34}$$

subject to the zero Dirichlet boundary condition:

$$u(x, y) = 0 \quad \text{when} \quad x = \pm \frac{s}{2} \quad \text{or} \quad y = \pm \frac{s}{2}. \quad (2.35)$$

The assumption $k = \text{const}$ is not necessary, and is introduced in equation (2.34) only for convenience.

Discretization of the Dirichlet boundary condition (2.35) is completely straightforward. Consider a Cartesian grid on the square Ω :

$$\left\{ (x_m, y_n) \mid m = -\frac{M}{2}, \dots, \frac{M}{2}, \quad n = -\frac{M}{2}, \dots, \frac{M}{2} \right\}, \quad (2.36)$$

$$M = s/h, \quad x_m = m \cdot h, \quad y_n = n \cdot h.$$

Since the scheme (2.32) is built on a compact 3×3 stencil, it does not require any additional “numerical” boundary conditions, and we simply discretize (2.35) as

$$u_{m,n} = 0 \quad \text{if} \quad m = \pm \frac{M}{2} \quad \text{or} \quad n = \pm \frac{M}{2},$$

which is exact and therefore accurate to any order.

It is known, however, that when solving the Helmholtz equation on a bounded domain subject to a Dirichlet boundary condition, resonances may occur. To avoid this undesirable phenomenon, we employ additional considerations when choosing the wavenumber k . Let the variable coefficients $a = a(x, y)$ and $b = b(x, y)$ in equation (2.34) be smooth and bounded on $\bar{\Omega}$. In addition, we require that

$$\nu \stackrel{\text{def}}{=} \min \left\{ \min_{(x,y) \in \bar{\Omega}} a(x, y), \min_{(x,y) \in \bar{\Omega}} b(x, y) \right\} > 0. \quad (2.37)$$

Inequality (2.37) implies, in particular, that the operator

$$\mathbf{L}u \equiv -\frac{\partial}{\partial x} \left(a \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(b \frac{\partial u}{\partial y} \right), \quad (2.38)$$

subject to the same Dirichlet boundary condition (2.35), is self-adjoint and positive definite on the space $W_{2,0}^2(\Omega)^2$. To guarantee uniqueness of the solution u to problem (2.34) with boundary condition (2.35), we must ensure that k^2 is not an eigenvalue of the operator \mathbf{L} of (2.38), (2.35). This is done by estimating the smallest eigenvalue λ_{\min} of \mathbf{L} , i.e., its eigenvalue closest to zero,

²This space is a completion in the norm $W_2^2(\Omega)$ of the set of functions $C_0^2(\bar{\Omega}) \subset C^2(\bar{\Omega})$ that are twice continuously differentiable on $\bar{\Omega}$ and are equal to zero on $\partial\Omega$. For self-adjointness on $W_{2,0}^2(\Omega)$, in addition to (2.37) one also needs to require an upper bound on the coefficients a and b , and on absolute values of their first derivatives, see [35, Sections 145, 148, 149].

and then choosing k accordingly.

As shown in [35, Section 150], the following estimate holds for the smallest eigenvalue of \mathbf{L} :

$$\lambda_{\min} \geq \frac{\nu}{c_{\Omega}},$$

where ν is defined in formula (2.37), and c_{Ω} is the constant from the Friedrichs inequality,

$$\iint_{\Omega} u^2(x, y) dx dy \leq c_{\Omega} \iint_{\Omega} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy. \quad (2.39)$$

When Ω is a square of side length s , it is easy to prove (see [36, Section 115]) that $c_{\Omega} = s^2$ (the area of the square) in inequality (2.39). Consequently,

$$\lambda_{\min} \geq \frac{\nu}{s^2}. \quad (2.40)$$

Inequality (2.40) implies that choosing the wavenumber k so that

$$k^2 < \frac{\nu}{s^2} \quad (2.41)$$

is sufficient for avoiding the resonances (and thus having uniqueness), since then the sum $\lambda_{\min} - k^2$ will remain positive.

In fact, estimate (2.40) is conservative and can be sharpened. If, for example, \mathbf{L} is the negative Laplace operator so that $a \equiv b \equiv 1$ and $\nu = 1$, then the first eigenfunction is $v = \cos(\pi x/s) \cos(\pi y/s)$ and the minimum eigenvalue is $\lambda_{\min} = 2\pi^2/s^2$. Hence, for practical purposes we estimate the minimum eigenvalue of \mathbf{L} by merely replacing the coefficients a and b in (2.38) by their minimum value ν of (2.37). This leads to a weaker constraint on k instead of (2.41):

$$k^2 < \frac{2\pi^2\nu}{s^2}. \quad (2.42)$$

In the numerical experiments of Section 2.5.1, we make sure that inequality (2.42) holds. This is a sufficient but not necessary condition for the uniqueness of a solution.

2.4.2 Local Sommerfeld-type boundary conditions

In order to test the performance of the scheme for larger values of k , the constraint given by inequality (2.42) must be alleviated. A convenient way of doing that is to modify the boundary condition so that the problem is no longer self-adjoint: its spectrum becomes complex, hence no real value $k^2 \in \mathcal{R}$ can be an eigenvalue.

In [34] (corrected in [37]) Erlangga and Turkel derived a fourth order accurate scheme for a

simplified absorbing boundary condition:

$$\frac{\partial u}{\partial x} + i\beta u = 0.$$

Furthermore, they present computational evidence that if the interior scheme is fourth order accurate but the absorbing boundary condition is only second order accurate, then the global accuracy is second order.

Here, we set the local Sommerfeld-type boundary conditions on two opposite sides of the square Ω :

$$\frac{\partial u}{\partial x} + iku = 0 \quad \text{if } x = \frac{s}{2}, \quad (2.43a)$$

$$\frac{\partial u}{\partial x} - iku = 0 \quad \text{if } x = -\frac{s}{2}, \quad (2.43b)$$

and keep a homogeneous Dirichlet boundary condition on the other pair of opposite sides:

$$u(x, y) = 0 \quad \text{if } y = \pm \frac{s}{2}. \quad (2.43c)$$

We emphasize that we do not intend to simulate the actual radiation of waves toward infinity by means of boundary conditions (2.43a) and (2.43b). The problem is still solved on a bounded region (a square), and our goal is rather to avoid the eigenvalues in the interior.

First, we again set the exact discrete Dirichlet BCs at the top and bottom edges of the square:

$$u_{m,n} = 0 \quad \text{if } n = \pm \frac{M}{2}. \quad (2.44a)$$

Boundary conditions (2.43a) and (2.43b) are then approximated on the grid with fourth order accuracy using compact differencing. For convenience, they are set at half-nodes:

$$u_{x_{M-\frac{1}{2},n}} + iku_{M-\frac{1}{2},n} = 0, \quad (2.44b)$$

$$u_{x_{\frac{1}{2},n}} - iku_{\frac{1}{2},n} = 0. \quad (2.44c)$$

We will treat the left boundary of the square $x = -\frac{s}{2}$, and the analogous case of the right boundary $x = \frac{s}{2}$ will follow by symmetry. First, we approximate (2.44c) with second order accuracy as follows:

$$\begin{aligned} u_{x_{\frac{1}{2},n}} - iku_{\frac{1}{2},n} &= \frac{u_{1,n} - u_{0,n}}{h} - \frac{(h/2)^2}{6} u_{xxx_{\frac{1}{2},n}} \\ &\quad - ik \frac{u_{1,n} + u_{0,n}}{2} + ik \frac{(h/2)^2}{2} u_{xx_{\frac{1}{2},n}} + \mathcal{O}(h^4). \end{aligned} \quad (2.45)$$

In order to eliminate the $\mathcal{O}(h^2)$ terms in (2.45), it is sufficient to approximate the derivatives $u_{xxx\frac{1}{2},n}$ and $u_{xx\frac{1}{2},n}$ with second order accuracy. To do so on the compact stencil, we will use the same difference formulae (2.27a) and (2.27b) obtained from the equation-based expressions in Section 2.3. Taking into account that $f(x, y)$ is compactly supported inside the square, we can assume that near the boundary $f \equiv 0$. For the derivative $u_{xx\frac{1}{2},n}$, this yields:

$$\begin{aligned} u_{xx\frac{1}{2},n} &= \frac{1}{a} (F_x - a_x u_x) \Big|_{\frac{1}{2},n} \\ &= \frac{1}{a_{\frac{1}{2},n}} \left(- (k^2 u)_{\frac{1}{2},n} - b_{y\frac{1}{2},n} \frac{u_{\frac{1}{2},n+1} - u_{\frac{1}{2},n-1}}{h} \right. \\ &\quad \left. - b_{\frac{1}{2},n} \frac{u_{\frac{1}{2},n+1} - 2u_{\frac{1}{2},n} + u_{\frac{1}{2},n-1}}{(h/2)^2} - a_{x\frac{1}{2},n} \frac{u_{1,n} - u_{0,n}}{h} \right) + \mathcal{O}(h^2). \end{aligned} \quad (2.46)$$

Similarly, for $u_{xxx\frac{1}{2},n}$ we obtain:

$$\begin{aligned} u_{xxx\frac{1}{2},n} &= \frac{1}{a} \left(\frac{\partial F_x}{\partial x} - a_{xx} u_x - \frac{2a_x}{a} (F_x - a_x u_x) \right) \Big|_{\frac{1}{2},n} \\ &= \frac{1}{a_{\frac{1}{2},n}} \left(- \left[\frac{b_{xy\frac{1}{2},n}}{h} (u_{\frac{1}{2},n+1} - u_{\frac{1}{2},n-1}) \right. \right. \\ &\quad + \frac{b_{y\frac{1}{2},n}}{4(h/2)^2} (u_{1,n+1} - u_{0,n+1} - u_{1,n-1} + u_{0,n-1}) \\ &\quad + \frac{b_{x\frac{1}{2},n}}{(h/2)^2} (u_{\frac{1}{2},n+1} + u_{\frac{1}{2},n-1} - 2u_{\frac{1}{2},n}) \\ &\quad + \frac{b_{\frac{1}{2},n}}{2(h/2)^3} (u_{1,n+1} - u_{0,n+1} + u_{1,n-1} - u_{0,n-1} \\ &\quad \left. \left. - 2(u_{1,n} - u_{0,n})) + \frac{(k^2 u)_{1,n} - (k^2 u)_{0,n}}{h} \right] \right. \\ &\quad - a_{xx\frac{1}{2},n} \frac{u_{1,n} - u_{0,n}}{h} - \frac{2a_{x\frac{1}{2},n}}{a_{\frac{1}{2},n}} \left(- (k^2 u)_{\frac{1}{2},n} - b_{y\frac{1}{2},n} \frac{u_{\frac{1}{2},n+1} - u_{\frac{1}{2},n-1}}{h} \right. \\ &\quad \left. \left. - b_{\frac{1}{2},n} \frac{u_{\frac{1}{2},n+1} - 2u_{\frac{1}{2},n} + u_{\frac{1}{2},n-1}}{(h/2)^2} - a_{x\frac{1}{2},n} \frac{u_{1,n} - u_{0,n}}{h} \right) \right) + \mathcal{O}(h^2). \end{aligned} \quad (2.47)$$

Formulae (2.46) and (2.47) still contain the values of u at semi-integer grid locations. To have a scheme that would only operate with full-node values, we replace $u_{\frac{1}{2},n}$ by the second order approximation $u_{\frac{1}{2},n} \approx \frac{u_{1,n} + u_{0,n}}{2}$, and proceed similarly for the terms $u_{\frac{1}{2},n+1}$, $u_{\frac{1}{2},n-1}$, and $(k^2 u)_{\frac{1}{2},n}$. Using this modification, we obtain a fourth order accurate approximation of the

boundary condition (2.43b) in the form

$$\left(\frac{u_{1,n} - u_{0,n}}{h} - \frac{h^2}{24} u_{xxx \frac{1}{2},n} \right) - ik \left(\frac{u_{1,n} + u_{0,n}}{2} - \frac{h^2}{8} u_{xx \frac{1}{2},n} \right) = 0, \quad (2.48)$$

where the terms $u_{xx \frac{1}{2},n}$ and $u_{xxx \frac{1}{2},n}$ are evaluated according to (2.46) and (2.47), respectively. Similarly, the boundary condition (2.43a) is approximated as

$$\left(\frac{u_{M,n} - u_{M-1,n}}{h} - \frac{h^2}{24} u_{xxx M-\frac{1}{2},n} \right) + ik \left(\frac{u_{M,n} + u_{M-1,n}}{2} - \frac{h^2}{8} u_{xx M-\frac{1}{2},n} \right) = 0, \quad (2.49)$$

where the terms $u_{xx M-\frac{1}{2},n}$ and $u_{xxx M-\frac{1}{2},n}$ are evaluated according to the analogues of (2.46) and (2.47), respectively.

For the case of the constant-coefficient Helmholtz equation dealt with in Chapters 3 and 4, the fourth-order formulae for the Sommerfeld BC (2.44) with the discrete representations (2.46-2.47) for the derivatives simplify as follows:

$$\begin{aligned} & \left(\frac{u_{M,n} - u_{M-1,n}}{h} - \frac{1}{6h} (u_{M,n+1} - u_{M-1,n+1} + u_{M,n-1} - u_{M-1,n-1} - 2(u_{M,n} - u_{M-1,n}) \right. \\ & \left. - \frac{k^2 h}{24} (u_{M,n} - u_{M-1,n}) \right) + ik \left(\frac{u_{M,n} + u_{M-1,n}}{2} + \frac{h^2 k^2}{8} u_{M-\frac{1}{2},n} \right. \\ & \left. + \frac{u_{M-\frac{1}{2},n+1} - 2u_{M-\frac{1}{2},n-1} + u_{M-\frac{1}{2},n-1}}{2} \right) = 0, \end{aligned} \quad (2.50a)$$

$$\begin{aligned} & \left(\frac{u_{1,n} - u_{0,n}}{h} - \frac{1}{6h} (u_{1,n+1} - u_{0,n+1} + u_{1,n-1} - u_{0,n-1} - 2(u_{1,n} - u_{0,n}) \right. \\ & \left. - \frac{k^2 h}{24} (u_{1,n} - u_{0,n}) \right) - ik \left(\frac{u_{1,n} + u_{0,n}}{2} + \frac{h^2 k^2}{8} u_{\frac{1}{2},n} + \frac{u_{\frac{1}{2},n+1} - 2u_{\frac{1}{2},n-1} + u_{\frac{1}{2},n-1}}{2} \right) = 0. \end{aligned} \quad (2.50b)$$

2.5 Numerical verification of the Cartesian scheme

To achieve the desired fourth order of accuracy in our compact finite difference approximation, the test solutions $u = u(x, y)$ for equation (2.34) must be at least 6 times continuously differentiable. Additionally, to satisfy the boundary condition (2.35) or boundary conditions (2.43), it is also convenient to choose the solution to be compactly supported inside the square D . We have found it easiest to devise such test solutions using polar coordinates, and we then convert back to Cartesian coordinates for actual computations. Specifically, we take a smooth and compactly

supported function of r and multiply it by a smooth 2π periodic function of θ for generality. In doing so, to guarantee regularity at $r = 0$, we additionally require that the function of r be equal to zero at the origin along with sufficiently many of its derivatives. Then, we substitute the resulting $u(r, \theta) = u(\sqrt{x^2 + y^2}, \arctan(y/x))$ into the left-hand side of equation (2.34) and derive the right-hand side $f = f(x, y)$, which is subsequently used in the scheme.

Our implementation was written in MATLAB, and the linear system obtained from our scheme is solved via MATLAB's built-in direct sparse solver. The computations were performed on a 2.16 GHz Intel Core 2 Duo MacBook Pro with 2 Gb of RAM running on Mac OS X.

The results in the following examples demonstrate fourth order convergence with respect to the grid size, and a somewhat faster than linear growth of the time required to compute the solution.

2.5.1 Dirichlet boundary condition

Example 1

For our first example, we use a test solution u , based on a trigonometric function of r , and coefficients a and b as follows:

$$\begin{aligned} u(x, y) &= \begin{cases} \sin^6(2r) \cos(\theta), & r < \frac{\pi}{2} \\ 0, & r > \frac{\pi}{2} \end{cases} \\ &= \begin{cases} \sin^6(2\sqrt{x^2 + y^2}) \cos(\arctan(y/x)), & r < \frac{\pi}{2} \\ 0, & r > \frac{\pi}{2} \end{cases}, \\ a(x, y) &= 1 + e^{-x^2 - y^2}, \\ b(x, y) &= 1 + \frac{x^2 + y^2}{1 + x^2 + y^2}. \end{aligned}$$

The domain in this case is a square of side length $s = \pi$ centered at the origin. The value of k was chosen to be $k = 1$. Note that in this case

$$\nu = \min_{(x, y) \in \bar{D}} \{a(x, y), b(x, y)\} = \min\{1 + e^{-2\pi^2}, 1 + 0\} = 1,$$

and thus we see that (2.42) is satisfied since

$$k^2 = 1 < 2 = \frac{2\pi^2}{\pi^2} = \frac{2\pi^2\nu}{s^2}.$$

Table 2.1 compares the error of the numerical and exact solutions on a series of grids of step-sizes h given in the leftmost column. From column 3 we clearly see the fourth order convergence.

Column 5 indicates that computational complexity of the direct solver scales somewhat faster than linear as the grid dimension increases.

Table 2.1: Grid convergence and computation time data for example 1.

h	$\ u - u_{num}\ _\infty$	Convergence Rate	Time	Time Scaling
1/4	$3.63 \cdot 10^{-2}$	-	$3.42 \cdot 10^{-3}$	-
1/8	$1.82 \cdot 10^{-3}$	4.47	$1.50 \cdot 10^{-2}$	4.38
1/16	$1.10 \cdot 10^{-4}$	4.07	$6.65 \cdot 10^{-2}$	4.44
1/32	$6.79 \cdot 10^{-6}$	4.02	.31	4.64
1/64	$4.22 \cdot 10^{-7}$	4.01	1.54	4.98
1/128	$2.64 \cdot 10^{-8}$	4.00	8.71	5.67

Example 2

In this example, we use a test solution u which is derived from a polynomial in r , with a and b as follows:

$$\begin{aligned}
u(x, y) &= \begin{cases} r^6(1 - r^2)^6 \sin(\theta), & r < 1 \\ 0, & r > 1 \end{cases} \\
&= \begin{cases} (x^2 + y^2)^3(1 - x^2 - y^2)^6 \sin(\arctan(y/x)), & r < 1 \\ 0, & r > 1 \end{cases}, \\
a(x, y) &= 1 + \frac{\arctan(xy)}{2\pi}, \\
b(x, y) &= 1 + \frac{x^2 + y^2}{1 + x^2 + y^2}.
\end{aligned}$$

Note that a is an asymmetric function of x and y , whereas both a and b were radially symmetric in the first example. The domain is a square of side length $s = 2$, centered at the origin. We choose $k = 1$, and so (2.42) is satisfied since

$$\nu = \min_{(x,y) \in \bar{D}} \{a(x, y), b(x, y)\} = \min\{1 + \frac{\arctan(-1)}{2\pi}, 1 + 0\} = \frac{7}{8}$$

and

$$k^2 = 1 < \frac{14\pi^2}{32} = \frac{2\pi^2\nu}{s^2}.$$

The computational results are summarized in Table 2.2. Column 3 clearly shows the fourth order convergence.

Table 2.2: Grid convergence and computation time data for example 2.

h	$\ u - u_{num}\ _\infty$	Convergence Rate	Time(s)	Time Scaling
1/4	$1.48 \cdot 10^{-3}$	-	$1.02 \cdot 10^{-3}$	-
1/8	$6.03 \cdot 10^{-5}$	4.96	$4.88 \cdot 10^{-3}$	4.77
1/16	$2.27 \cdot 10^{-6}$	5.16	$2.10 \cdot 10^{-2}$	4.30
1/32	$1.24 \cdot 10^{-7}$	4.28	0.11	5.05
1/64	$7.42 \cdot 10^{-9}$	4.09	0.56	5.30
1/128	$4.63 \cdot 10^{-10}$	4.00	3.23	5.76

Example 3

We now use a test solution u which includes an exponential function in r :

$$\begin{aligned}
u(x, y) &= \begin{cases} (1 - r^2)^6 (1 - e^{-r^2})^6 \sin(\cos(\theta)), & r < 1 \\ 0, & r > 1 \end{cases} \\
&= \begin{cases} (1 - x^2 - y^2)^6 (1 - e^{-x^2 - y^2})^6 \sin(\cos(\arctan(y/x))), & r < 1 \\ 0, & r > 1 \end{cases}, \\
a(x, y) &= 1 + \frac{\arctan(x + y)}{2\pi}, \\
b(x, y) &= 1 + \frac{e^{-x^2 - y^2}}{1 + e^{-x^2 - y^2}}.
\end{aligned}$$

The domain is a square of side length $s = 2$ centered at the origin. We choose $k = 1$ and verify that (2.42) is satisfied:

$$\nu = \min_{(x,y) \in D} \{a(x, y), b(x, y)\} = \min\left\{1 + \frac{\arctan(-2)}{2\pi}, 1\right\} \approx 0.8238.$$

We can then compute the bound in (2.42) to be approximately 3.3489, and with our choice of $k=1$ we indeed see that (2.42) is satisfied. Table 2.3 summarizes the numerical results. As before column 3 demonstrates the fourth order convergence, and column 5 shows that computational complexity scales faster than linear as the grid dimension increases.

Table 2.3: Grid convergence and computation time data for example 3.

h	$\ u - u_{num}\ _{\infty}$	Convergence Rate	Time(s)	Time Scaling
1/4	$5.13 \cdot 10^{-5}$	-	$1.30 \cdot 10^{-3}$	-
1/8	$2.71 \cdot 10^{-6}$	4.36	$4.94 \cdot 10^{-3}$	3.79
1/16	$1.07 \cdot 10^{-7}$	5.03	$2.06 \cdot 10^{-2}$	4.18
1/32	$5.43 \cdot 10^{-9}$	4.43	0.11	5.24
1/64	$3.25 \cdot 10^{-10}$	4.09	0.56	5.14
1/128	$2.016 \cdot 10^{-11}$	4.01	3.24	5.82

2.5.2 Local Sommerfeld-type boundary conditions

To demonstrate the effectiveness of using the Sommerfeld-type boundary condition to eliminate resonances, we now let $k = 20$ in each of the three previous examples of Section 2.5.1, and use the Sommerfeld-type boundary conditions (2.43) instead of the Dirichlet boundary condition (2.35). In all of the following examples, we take a square domain of side length $s = 4$.

Example 4

Table 2.4: Grid convergence and computation time data for example 4.

h	$\ u - u_{num}\ _{\infty}$	Convergence Rate	Time(s)	Time Scaling
1/8	$2.01 \cdot 10^{-3}$	-	$3.41 \cdot 10^{-2}$	-
1/16	$1.07 \cdot 10^{-4}$	4.34	0.19	5.69
1/32	$6.35 \cdot 10^{-6}$	4.10	1.07	5.51
1/64	$4.01 \cdot 10^{-7}$	3.98	7.20	6.74
1/128	$2.51 \cdot 10^{-8}$	4.00	48.74	6.77

Example 5

Table 2.5: Grid convergence and computation time data for example 5.

h	$\ u - u_{num}\ _{\infty}$	Convergence Rate	Time(s)	Time Scaling
1/8	$4.69 \cdot 10^{-4}$	-	$2.94 \cdot 10^{-2}$	-
1/16	$1.54 \cdot 10^{-5}$	5.52	0.20	6.80
1/32	$8.98 \cdot 10^{-7}$	4.14	1.15	5.77
1/64	$4.84 \cdot 10^{-8}$	4.31	8.39	7.27
1/128	$2.71 \cdot 10^{-9}$	4.23	53.00	6.31

Example 6

Table 2.6: Grid convergence and computation time data for example 6.

h	$\ u - u_{num}\ _{\infty}$	Convergence Rate	Time(s)	Time Scaling
1/8	$1.63 \cdot 10^{-5}$	-	$3.49 \cdot 10^{-2}$	-
1/16	$3.77 \cdot 10^{-7}$	6.57	0.22	6.20
1/32	$2.35 \cdot 10^{-8}$	4.00	1.21	5.58
1/64	$1.45 \cdot 10^{-9}$	4.02	7.50	6.21
1/128	$8.92 \cdot 10^{-11}$	4.03	59.10	7.88

2.6 Auxiliary problems for the method of difference potentials

In this chapter, we have demonstrated the equation-based approach for building high-order FD schemes in 1D and 2D. While schemes for the Helmholtz equation in polar coordinates in both one and two dimensions were derived, the remaining chapters will make use only of the Cartesian scheme (2.33) of Section 2.3, even though the domain used in those chapters is a disk centered at the origin. This will demonstrate that the method of difference potentials in

Chapter 3 allows us to apply this Cartesian scheme on domains for which the boundary curve is not aligned with the grid without any loss of accuracy. As such, one and the same Cartesian finite difference scheme can be used for a variety of nonconforming boundary shapes.

In order to apply the method of difference potentials, we will embed the domain Ω in a larger domain Ω_0 . The larger domain Ω_0 will be used to formulate what is known as the auxiliary problem (AP). When formulating the AP, the only essential requirement is the existence and uniqueness of its solution on Ω_0 for any right-hand side g , and, of course, well-posedness (i.e., the continuous dependence of the solution on the data). Otherwise, the choice of AP will not affect the solution that we ultimately obtain inside Ω . It is to our advantage then to formulate an AP that is easily solvable, and this is precisely the motivation behind our choice of the square domain, as well as of the boundary conditions in (2.43). Moreover, this explains why we have concerned ourselves in this chapter with only problems that have their support on the interior of the square domain (which permits us to implement the Sommerfeld-type BC to avoid resonances³), since the domain of interest, Ω , will be strictly in the interior of the square auxiliary domain Ω_0 with side-length s . Altogether, the continuous AP is given by:

$$\begin{aligned} \mathbf{L}u &= g, & \mathbf{x} &\in \Omega_0, \\ u &= 0, & y &= \pm \frac{s}{2}, \\ \frac{\partial u}{\partial x} \pm iku &= 0, & x &= \pm \frac{s}{2}, \end{aligned} \tag{2.51}$$

where the right-hand side g is an auxiliary right-hand side that results from the method of difference potentials rather than the physical source term of the problem.

Indeed, the AP (2.51) can be solved efficiently by means of the separation of variables⁴. As pointed out in Section 2.4.2, it is known that the Helmholtz equation is prone to resonances if only Dirichlet boundary conditions are used,⁵ whereas the Sommerfeld-type conditions of (2.51) make the spectrum complex and hence guarantee uniqueness of the solution to the AP. The disadvantage of using Sommerfeld type conditions is that they introduce complex quantities into the calculation, which is not always necessary for interior problems (unlike for the exterior problems). Alternatively, one could use a Dirichlet or real Robin condition that was carefully chosen to avoid zero (or very small) eigenvalues and would keep the solution real, and this approach was adopted for some of the computations in [18].

³If, however, one is interested in solving problems on rectangular domains with other boundary conditions than we have described, then Section 2.4.2 serves as an example of how to use the equation-based approach to build high order approximations to the BCs on the compact stencil.

⁴Separation of variables can only be done in the case of the constant-coefficient Helmholtz equation, which we use throughout Chapters 3 and 4. In the special case that $a = a(x)$ and $b = b(y)$ separation of variables may also be used, but an LU-solver is employed for more general variable-coefficient equations.

⁵The Helmholtz equation is said to be at a resonance on Ω_0 if $-k^2$ is an eigenvalue of the Laplacian subject to zero boundary conditions at $\partial\Omega_0$. In this case, the solution to the Helmholtz equation is not unique.

To form the discrete problem, we apply the compact scheme (2.33) to the differential equation $\mathbf{L}u = g$ of (2.51):

$$\begin{aligned}
& \frac{1}{h^2} (u_{m+1,n} + u_{m,n+1} + u_{m-1,n} + u_{m,n-1} - 4u_{m,n}) \\
& + \frac{1}{6h^2} [u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} + 4u_{m,n} \\
& - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})] \\
& + \frac{k^2}{12} (u_{m+1,n} + u_{m,n+1} + 8u_{m,n} + u_{m-1,n} + u_{m,n-1}) = g_{m,n}^{(h)},
\end{aligned} \tag{2.52}$$

where we can formally think that [cf. formula (2.33)]

$$g_{m,n}^{(h)} = g_{m,n} + \frac{1}{12} (g_{m+1,n} + g_{m,n+1} - 4g_{m,n} + g_{m-1,n} + g_{m,n-1}).$$

We emphasize, however, that the right-hand side g in (2.51) plays only an auxiliary role: it will be present even if the governing equation is homogeneous, and even in the inhomogeneous case g is related but not identical to the source term of the PDE. As such, in our subsequent analysis the explicit form of $g_{m,n}$ will never be needed. What will rather be important for constructing the difference potentials and projections is the final discrete right-hand side $g_{m,n}^{(h)}$ of equation (2.52). This right-hand side $g^{(h)}$ will be obtained directly, i.e., without having to relate it to any g from the continuous AP (2.51) by means of the five-node stencil. The auxiliary right-hand side $g^{(h)}$ will be specified in Chapter 3 by the definition of the difference potential (3.2). We also note that even though we keep the right-hand side $g^{(h)}$ as a key innate element of the method of difference potentials, the actual physical solutions that we obtain inside Ω in the form of difference potentials will be those to the finite difference equation (2.33). Additionally, the BCs of (2.51) are discretized according to (2.44) with the fourth-order approximations (2.50) for the Sommerfeld-type BCs for the constant-coefficient Helmholtz equation.

As in the case of the continuous AP (2.51), the overall discrete AP (2.52) is supposed to have a unique solution $u_{m,n}$, $m = 0, \dots, M$, $n = 0, \dots, N$, for any right-hand side $g_{m,n}$ (defined on the interior sub-grid $m = 1, \dots, M-1$, $n = 1, \dots, N-1$), and be well-posed. For the constant-coefficient Helmholtz equation, the discrete AP (2.52) can be solved by a sine FFT in the y -direction combined with the tri-diagonal elimination in the x -direction. The complexity of this solution is log-linear with respect to the grid dimension N , and linear with respect to M (note here that we have chosen a square domain, and, consequently, $M = N$).

In this way, the Cartesian scheme obtained in Section 2.3 with the boundary conditions (2.43) will be used to solve the auxiliary problem in the method of difference potentials in Chapters 3 and 4. Ultimately, this will yield the solution to the Helmholtz equation on a non-Cartesian domain Ω without the usually associated loss of accuracy due to staircasing.

Chapter 3

Difference Potentials

In this chapter, we introduce the method of difference potentials and present a numerical approach for its universal treatment of boundary conditions [25]. In Section 3.1 we discuss difference potentials and projections, which can be considered as generalized discrete counterparts of Calderon’s potentials and projections [6, 7]. We show how the finite difference equations on the domain of interest can be reduced to equivalent discrete equations at the boundary after encapsulating it in a larger auxiliary domain. Our main objective is to demonstrate the versatility of the proposed algorithm in treating boundary conditions. Hence, we consider a computational setting which is otherwise very straightforward. Specifically, we use the fourth-order accurate compact scheme of Section 2.3 to solve the inhomogeneous interior problem for the Helmholtz equation (1.1) on a uniform Cartesian grid in two space dimensions. Moreover, we simplify the FD scheme to the case of constant coefficients. For the preliminary presentation of the method, the computational domain is chosen as a disk — the simplest non-conforming shape for our Cartesian finite difference grid. We emphasize that these restrictions are not limitations of our approach, and this is done only to isolate the discussion of complicated boundary conditions in an otherwise simple setting.

In Section 3.2, we provide a detailed account of how various types of boundary conditions can be accommodated by the method of difference potentials. We illustrate the generality of this approach by analyzing specific examples: Robin boundary conditions with variable coefficients and mixed Dirichlet/Neumann boundary conditions. We identify those parts of the overall numerical algorithm that need to be adjusted when changing the boundary conditions and show that only minor changes are needed in the algorithm to accommodate changes in the boundary condition.

In Section 3.3, we present the results of the numerical experiments. Our simulations corroborate the theoretical design properties of the algorithm. Specifically, when the solution is sufficiently smooth, the algorithm demonstrates fourth-order grid convergence. We also offer

simulations to demonstrate that the grid convergence predictably slows down when the solution lacks regularity due to singularities in the boundary data. Restoring the design convergence rate when the solution has singularities at the boundary is the topic of Chapter 4. Additionally, we show that boundary value problems with the same domain but different boundary conditions or a different inhomogeneous source term can be solved very efficiently with no special changes required in the algorithm.

3.1 Difference potentials and projections

Let Ω be a bounded domain on the Cartesian plane \mathbb{R}^2 , and let Γ be its boundary, $\Gamma = \partial\Omega$. Consider the following boundary value problem:

$$Lu \stackrel{\text{def}}{=} \Delta u + k^2 u = f, \quad \mathbf{x} \in \Omega, \quad (3.1a)$$

$$l_\Gamma u = \phi_\Gamma, \quad (3.1b)$$

where $k = \text{const}$ in equation (3.1a). Problem (3.1) is required to be well-posed, having a unique solution u on Ω for a given ϕ_Γ . We discretize problem (3.1) on a Cartesian grid and solve it with high-order accuracy using the method of difference potentials for the case where Ω is a disk of radius $r = 1$ centered at the origin and Γ is a circle. Note that (3.1b) is a generic boundary condition that will be specified later. The method allows for a broad variety of boundary conditions (3.1b) which will be explored in Section 3.2. Throughout the presentation of the method, the source term f of (3.1a) is always decoupled from the expressions resulting from the differential operator, allowing different source terms or homogeneous problems to be treated with ease.

3.1.1 The finite difference scheme and auxiliary problem

The method of difference potentials can be applied in conjunction with any finite difference scheme. A key advantage of high-order schemes is their improved efficiency in reducing the phase error, see Section 1. Hereafter, we restrict the discussion to the constant coefficient case, since our focus is on the treatment of the boundary conditions. In the case of constant coefficients or even a variable wavenumber, a sixth order accurate FD scheme is possible [31, 32], but we have chosen to implement the fourth-order accurate approximation (2.33) of Section 2.3 for the inhomogeneous constant-coefficient Helmholtz equation (3.1a).

In order to apply the method of difference potentials, we formulate an auxiliary problem on a larger square domain Ω_0 which includes Ω (i.e., the unit disk) as described in Section 2.6. For the disk of radius 1 centered at the origin, we choose Ω_0 to be a square of side length $s = 2.2$ (also centered at the origin). The continuous AP is posed as (2.51) with zero Dirichlet BCs on the top

and bottom edges of the square and Sommerfeld-type BCs on the left and right edges to avoid resonances of the Helmholtz equation. The discrete AP subject to the finite difference scheme (2.33) for the constant-coefficient Helmholtz equation is given by (2.52), with the discrete BCs given in (2.44) computed by the formulae (2.50). This discrete auxiliary problem can be solved efficiently (see the concluding remarks of Section 2.6). The method of difference potentials will require that we solve several such APs for a single problem, as will be made clear in the forthcoming sections.

We now consider a Cartesian grid on the square auxiliary domain in the next section. We will define various grid sets that are useful for building the discrete operators of the method of difference potentials.

3.1.2 Grid sets and operators

Grid subsets Let \mathbb{N}_0 be a uniform Cartesian grid on the square Ω_0 with step size h in both the x - and y -directions, and let $\mathbb{M}_0 \subset \mathbb{N}_0$ be the set of its interior nodes, i.e., all nodes of \mathbb{N}_0 except those on the edges of Ω_0 . The solution u to the discrete auxiliary problem (2.52) will be defined on the grid \mathbb{N}_0 , while its right-hand side $g^{(h)}$ will be defined on the grid \mathbb{M}_0 .

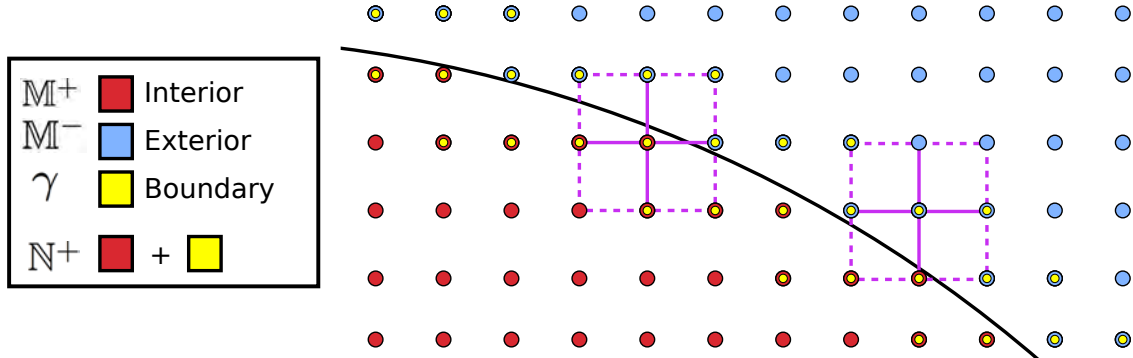


Figure 3.1: Interior and exterior grid subsets and the grid boundary.

Let $\mathbb{M}^+ \subset \mathbb{M}_0$ be the set of nodes of \mathbb{M}_0 that are inside the physical domain Ω , i.e., $\mathbb{M}^+ = \mathbb{M}_0 \cap \Omega$, see the red nodes of Figure 3.1. Since $\Gamma = \partial\Omega$ is not aligned with the grid, we define its discrete analogue, γ , which we refer to as the grid boundary. Let \mathbb{M}^- be the set of all nodes of \mathbb{M}_0 that lie outside Ω , i.e., $\mathbb{M}^- = \mathbb{M}_0 \setminus \mathbb{M}^+$, see the blue nodes of Figure 3.1. Let \mathbb{N}^+ and \mathbb{N}^- be defined as the sets of nodes of \mathbb{N}_0 that are used when applying the 3×3 compact stencil (see Figure 2.1) to the nodes of \mathbb{M}^+ and \mathbb{M}^- , respectively. In Figure 3.1, \mathbb{N}^+ consists of

the red plus yellow nodes, while N^- is denoted by the blue plus yellow nodes. By design, there will be an overlap in the sets N^+ and N^- when applying the stencil to nodes of M^+ and M^- which are adjacent to Γ , and we refer to this overlap, i.e., the intersection of these sets, as the grid boundary, $\gamma = N^+ \cap N^-$, see the yellow nodes of Figure 3.1.

Difference potentials The solution of the discrete AP (2.52) plays a key role in the construction of the difference potentials and projections, which can be considered as discrete counterparts of Calderon's potentials and boundary projections (pseudo-differential operators), see [6, 7, 8]. The difference potential will approximate the solution u of boundary value problem (3.1) on the grid N^+ . The density of the difference potential is a grid function defined on the grid boundary γ . It satisfies a special system of linear algebraic equations called the boundary equation with projection (BEP).

We denote the discrete operator on the left-hand side of equation (2.52) by $L^{(h)}$. Let us first discuss the operators for the homogeneous Helmholtz equation with $f = 0$. Then the discrete AP (2.52) consists of solving the finite difference equation $L^{(h)}u = g^{(h)}$ on the grid N_0 , subject to boundary conditions (2.44) computed by formulae (2.50a) and (2.50b). Define the corresponding inverse operator $G^{(h)}$ as the solution of the discrete AP (2.52), so that $u = G^{(h)}g^{(h)}$. Also consider a grid function ξ_γ specified on the discrete boundary γ , Figure 3.1. The difference potential with density ξ_γ is given by

$$P_{N^+}\xi_\gamma \stackrel{\text{def}}{=} w - \underbrace{G^{(h)}(L^{(h)}w|_{M^+})}_{g^{(h)}}, \quad \text{where} \quad w = \begin{cases} \xi_\gamma & \text{on } \gamma, \\ 0 & \text{on } N_0 \setminus \gamma. \end{cases} \quad (3.2)$$

The operation $L^{(h)}w|_{M^+} := g^{(h)}$ in formula (3.2) denotes first the application of the operator $L^{(h)}$ to the auxiliary function w and then truncation of the grid function $L^{(h)}w$ to the grid M^+ , see Figure 3.1. Note also that $g^{(h)} := L^{(h)}w|_{M^+}$ is the directly obtained right-hand side of the discrete AP (2.52). As mentioned there is no explicit function g from the definition of the continuous AP (2.51) to which we apply the 5-node stencil implied by the finite difference scheme (2.33). The difference potential $P_{N^+}\xi_\gamma$ is defined on the grid N^+ (that's why we are using the subscript " N^+ "),¹ and at the nodes M^+ it satisfies the finite difference equation [cf. formula (2.33)]

$$L^{(h)}(P_{N^+}\xi_\gamma) = 0.$$

Along with the grid function ξ_γ , consider a two-component vector function $\xi_\Gamma = (\xi_0, \xi_1)$

¹Even though both the auxiliary function w and the solution $G^{(h)}g$ to the discrete AP (2.52) are defined on the entire grid N_0 , the difference potential $P_{N^+}\xi_\gamma$ as introduced by formula (3.2) is of interest only on the grid N^+ .

defined on the continuous boundary Γ . ξ_Γ can be used as the density of the genuine Calderon potential of the differential operator \mathbf{L} on the domain Ω [6, 7], in which ξ_0 is interpreted as the Dirichlet data and ξ_1 is interpreted as the Neumann data at the boundary Γ . Provided that the grid function ξ_γ is related to the continuous function ξ_Γ in some special sense (ξ_γ must be obtained from ξ_Γ as an equation-based extension based on the Taylor formula of order four, see Section 3.2), the difference potential (3.2) approximates the continuous Calderon potential with density ξ_Γ with fourth-order accuracy on the grid \mathbb{N}^+ (the design accuracy of our compact scheme), see [8, 38, 18].

Truncating the difference potential (3.2) to γ , we obtain the difference projection:

$$\mathbf{P}_\gamma \xi_\gamma \stackrel{\text{def}}{=} (\mathbf{P}_{\mathbb{N}^+} \xi_\gamma)|_\gamma, \quad (3.3)$$

and then express the discrete BEP for the homogeneous problem as:

$$\mathbf{P}_\gamma \xi_\gamma = \xi_\gamma. \quad (3.4)$$

Its pivotal property (see [8]) is that a grid function ξ_γ satisfies the BEP (3.4) if and only if it can be obtained as the truncation to γ of a solution u [defined on \mathbb{N}^+ , Figure 3.1] of the homogeneous difference equation (2.33) with $f = 0$: $\mathbf{L}^{(h)}u = 0$. Thus, the BEP (3.4) provides an equivalent reduction of the discrete equation $\mathbf{L}^{(h)}u = 0$ from the grid domain \mathbb{N}^+ to the grid boundary γ . If the grid function u satisfies $\mathbf{L}^{(h)}u = 0$, then its truncation $\xi_\gamma = u|_\gamma$ must satisfy the BEP (3.4). Conversely, if the grid function ξ_γ satisfies the BEP (3.4), then there exists a function u defined on \mathbb{N}^+ such that $\mathbf{L}^{(h)}u = 0$ and $u|_\gamma = \xi_\gamma$. In fact, this u is given by the difference potential (3.2): $u = \mathbf{P}_{\mathbb{N}^+} \xi_\gamma$.

In Chapter 4, we will need the ability to solve inhomogeneous problems since singular problems, even when homogeneous, require the solution of an inhomogeneous “regularized” problem when using the method of singularity subtraction presented there. When the Helmholtz equation (1.1) is inhomogeneous (i.e., the right-hand side f is nonzero), the BEP (3.4) also becomes inhomogeneous:

$$\mathbf{P}_\gamma \xi_\gamma + \mathbf{Tr}_\gamma \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)} = \xi_\gamma. \quad (3.5)$$

The function $f^{(h)}$ is equal to the right-hand side f on the interior nodes of the discrete domain (since f is only defined inside of the given domain) and is equal to zero on the exterior nodes (see Figure 3.1). $\mathbf{B}^{(h)}$ denotes the 5-node stencil operator of the finite difference scheme (2.33), and the operator $\mathbf{Tr}_\gamma \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$ represents the trace on the grid boundary γ of a grid function $\mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$ defined on \mathbb{N}_0 . The extension $f^{(h)}$ of f is necessary since the operator $\mathbf{G}^{(h)}$ is defined by the auxiliary problem and operates on the entire auxiliary domain. Consequently, the difference potential from which we obtain the solution now also contains an inhomogeneous

term:

$$u = \mathbf{P}_{\mathbb{N}^+} \xi_\gamma + \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}. \quad (3.6)$$

Note that the difference potential (3.2) and, consequently, the projection (3.3) depends on the choice of the AP since changing the AP will change the inverse operator $\mathbf{G}^{(h)}$. However, the change of the AP does not affect the range of the projection \mathbf{P}_γ , i.e., it does not change the set of solutions to the BEP (3.4), as long as the AP remains uniquely solvable. In other words, when changing the AP one only induces a different projection onto the same subspace of solutions, see [8].

3.2 Treatment of the boundary conditions

In this section, we show how to account for the given boundary condition (3.1b) in order to approximately reconstruct the data $(u, \frac{\partial u}{\partial \mathbf{n}})|_\Gamma$ of the solution u to problem (3.1) at the boundary Γ , and subsequently obtain the discrete solution of (3.1) on the grid \mathbb{N}^+ in the form of the difference potential (3.2). It is from the treatment of the BCs by a truncated series expansion in Section 3.2.2 that the need for solving several APs (2.52) will arise. First, we need to build the equation-based extension of an arbitrary pair of functions, $\xi_\Gamma = (\xi_0, \xi_1)|_\Gamma$, from the continuous boundary Γ to the grid boundary γ in order to apply the discrete operators of Section 3.1.2.

3.2.1 Equation-based extension to the grid boundary

The extension of a given $\xi_\Gamma = (\xi_0, \xi_1)|_\Gamma$ from the smooth boundary Γ to the grid nodes γ (specifically the nodes adjacent to Γ , see Figure 3.1) is constructed using a truncated Taylor expansion with differentiation in the direction normal to Γ . Consider ξ_Γ as the data of some function $v = v(x, y)$,

$$(\xi_0, \xi_1)|_\Gamma = \left(v, \frac{\partial v}{\partial \mathbf{n}} \right) \Big|_\Gamma,$$

and we define the function v near the curve Γ by means of the Taylor expansion:

$$v \stackrel{\text{def}}{=} v_\Gamma + \rho \frac{\partial v}{\partial \mathbf{n}} \Big|_\Gamma + \frac{\rho^2}{2} \frac{\partial^2 v}{\partial \mathbf{n}^2} \Big|_\Gamma + \frac{\rho^3}{6} \frac{\partial^3 v}{\partial \mathbf{n}^3} \Big|_\Gamma + \frac{\rho^4}{24} \frac{\partial^4 v}{\partial \mathbf{n}^4} \Big|_\Gamma. \quad (3.7)$$

In formula (3.7), ρ denotes the distance (with sign) from a given point near Γ to the curve Γ . We emphasize that while formula (3.7) takes the usual form of a Taylor approximation to the function v , it should not be interpreted this way. Instead, it should be thought of as the definition of v . The new function v can be evaluated at any point (x, y) which is sufficiently close to Γ . In particular, we call this new function ξ_γ when its domain is restricted to the nodes

of the grid boundary γ , see Figure 3.1:

$$\xi_\gamma \stackrel{\text{def}}{=} v|_\gamma. \quad (3.8)$$

However, the foregoing definition is not complete until we specify how to compute the normal derivatives of order 2 and higher required for formula (3.7). These will be obtained using equation-based differentiation applied to the homogeneous Helmholtz equation with constant coefficients, (3.1a) with $f = 0$. It will be convenient to treat the homogeneous case first, since we will compute the homogeneous and inhomogeneous components of the extension separately. The inhomogeneous term will be obtained by the separating out the right-hand side f and its derivatives from the same process which we now apply to the homogeneous equation.

First, we assume that the “input” functions v and $\frac{\partial v}{\partial n}$ are known analytically on the circle Γ so that we can readily compute their tangential derivatives. Since we consider the simple case of a circular boundary Γ centered at the origin, the outward normal to Γ and the direction of the polar radius coincide. Hence, it is convenient to recast equation (3.1a) using polar coordinates (r, θ) :

$$\frac{1}{r} \frac{\partial v}{\partial r} + \frac{\partial^2 v}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + k^2 v = 0. \quad (3.9)$$

Equation (3.9) allows us to obtain the second derivative of v with respect to r :

$$\frac{\partial^2 v}{\partial r^2} = - \left(\frac{1}{r} \frac{\partial v}{\partial r} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + k^2 v \right). \quad (3.10)$$

Recall that v and $\frac{\partial v}{\partial n} = \frac{\partial v}{\partial r}$ are given on Γ , and that $\frac{\partial^2 v}{\partial \theta^2}$ can be computed analytically as the second tangential derivative of the given function v . Hence, (3.10) allows us to compute the term $\frac{\partial^2 v}{\partial n^2} = \frac{\partial^2 v}{\partial r^2}$ in the Taylor expansion (3.7).

We proceed to find the remaining normal derivatives, $\frac{\partial^3 v}{\partial n^3} = \frac{\partial^3 v}{\partial r^3}$ and $\frac{\partial^4 v}{\partial n^4} = \frac{\partial^4 v}{\partial r^4}$, via equation-based differentiation². In particular, we take the derivative of (3.10) with respect to r to obtain:

$$\frac{\partial^3 v}{\partial r^3} = \frac{1}{r^2} \frac{\partial v}{\partial r} - \frac{1}{r} \frac{\partial^2 v}{\partial r^2} + \frac{2}{r^3} \frac{\partial^2 v}{\partial \theta^2} - \frac{1}{r^2} \frac{\partial^3 v}{\partial r \partial \theta^2} - k^2 \frac{\partial v}{\partial r}. \quad (3.11)$$

We are able to evaluate $\frac{\partial^3 v}{\partial r^3}$ using the given function $\frac{\partial v}{\partial r}$, the analytically computed second tangential derivative of $\frac{\partial v}{\partial r}$ for $\frac{\partial^3 v}{\partial r \partial \theta^2}$, and the representation (3.10) for $\frac{\partial^2 v}{\partial r^2}$. This way, we can compute the third normal derivative term of the Taylor expansion (3.7). To compute the next

²This is the same method by which we obtained alternate representations of the truncation terms for finite difference schemes in Chapter 2.

term, we differentiate equation (3.11) with respect to r :

$$\frac{\partial^4 v}{\partial r^4} = -\frac{2}{r^3} \frac{\partial v}{\partial r} + \frac{2}{r^2} \frac{\partial^2 v}{\partial r^2} - \frac{1}{r} \frac{\partial^3 v}{\partial r^3} - \frac{6}{r^4} \frac{\partial^2 v}{\partial \theta^2} + \frac{4}{r^3} \frac{\partial^3 v}{\partial r \partial \theta^2} - \frac{1}{r^2} \frac{\partial^4 v}{\partial r^2 \partial \theta^2} - k^2 \frac{\partial^2 v}{\partial r^2}. \quad (3.12)$$

Again, the term $\frac{\partial v}{\partial r}$ is given, the terms $\frac{\partial^2 v}{\partial r^2}$ and $\frac{\partial^3 v}{\partial r^3}$ are evaluated via (3.10) and (3.11), respectively, and the terms $\frac{\partial^2 v}{\partial \theta^2}$ and $\frac{\partial^3 v}{\partial r \partial \theta^2}$ are computed analytically as the tangential derivatives of the given pair of functions $(v, \frac{\partial v}{\partial r})$. The only remaining term of (3.12) that has not been accounted for yet is $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$. We evaluate it by differentiating equation (3.10) twice with respect to θ :

$$\frac{\partial^4 v}{\partial r^2 \partial \theta^2} = - \left(\frac{1}{r} \frac{\partial^3 v}{\partial r \partial \theta^2} + \frac{1}{r^2} \frac{\partial^4 v}{\partial \theta^4} + k^2 \frac{\partial^2 v}{\partial \theta^2} \right). \quad (3.13)$$

Note that all of the terms required to compute $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$ by (3.13) are tangential (i.e., angular) derivatives of v and $\frac{\partial v}{\partial r}$. Therefore, substituting (3.13) for $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$ into (3.12) completes our ability to calculate the fourth normal derivative $\frac{\partial^4 v}{\partial r^4} = \frac{\partial^4 v}{\partial n^4}$ in the Taylor expansion (3.7).

Thus, given an arbitrary pair of functions $(v, \frac{\partial v}{\partial r})$ defined along the circle Γ , a fifth-order extension from the continuous circle to a nearby grid node along the normal direction is accomplished by substituting equations (3.10–3.13) into the Taylor expansion (3.7). In [18, Appendix A], a similar extension is built in the case of an arbitrary smooth curve Γ .

Hereafter, we use the notation \mathbf{Ex} for the equation-based extension operator defined by formulae (3.7–3.8). It will act on an arbitrary pair of continuous functions defined on Γ : $\xi_\Gamma = (\xi_0, \xi_1)|_\Gamma$. This operator uses the truncated Taylor expansion (3.7) to construct a new function $v(x, y)$ near Γ , which is then sampled at the grid boundary γ according to (3.8). This yields the grid function that we refer to as ξ_γ :

$$\xi_\gamma = \mathbf{Ex} \xi_\Gamma = \mathbf{Ex} (\xi_0, \xi_1) |_\gamma.$$

We emphasize that while the normal derivatives of order two and higher in formula (3.7) are obtained by differentiation based on the Helmholtz equation (3.1a), the construction of the operator \mathbf{Ex} permits it to be applied to an arbitrary pair of functions. Thus, ξ_Γ does not need to represent the data of a solution to equation (3.1a) in order to apply the operator. However, if ξ_Γ happens to be the data of a solution u to equation (3.1a) on Ω , then formula (3.7) approximates this solution near $\Gamma = \partial\Omega$ with fifth-order accuracy with respect to n , and, in particular, it shall do so at the nodes of the discrete boundary γ .

When the Helmholtz equation is inhomogeneous, which is unavoidable for problems with boundary singularities (see Chapter 4), this extension process requires some modification. In particular, the right-hand side of (3.9) becomes f rather than zero, and its differentiation must be carried through in the subsequent derivations (3.10–3.13). It is convenient to separate out the

contribution of the right-hand side f of (3.9) in the inhomogeneous case as it applies to formula (3.7), which equivalently amounts to decoupling expressions (3.10–3.13) into the homogeneous part and inhomogeneous part. In doing so, the Taylor expansion (3.7) is decomposed into the homogeneous and inhomogeneous contribution as follows:

$$\begin{aligned}
v = & \underbrace{v_\Gamma + \rho \frac{\partial v}{\partial r} \Big|_\Gamma + \frac{\rho^2}{2} \frac{\partial^2 v}{\partial r^2} \Big|_\Gamma + \frac{\rho^3}{6} \frac{\partial^3 v}{\partial r^3} \Big|_\Gamma + \frac{\rho^4}{24} \frac{\partial^4 v}{\partial r^4} \Big|_\Gamma}_{\text{homogeneous}} \\
& + \underbrace{\frac{\rho^2}{2} f + \frac{\rho^3}{6} \left(\frac{\partial f}{\partial r} - \frac{1}{r} f \right) + \frac{\rho^4}{24} \left(\frac{\partial^2 f}{\partial r^2} - \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} - \frac{1}{r} \frac{\partial f}{\partial r} - k^2 f + \frac{3}{r^2} f \right)}_{\text{inhomogeneous}}. \tag{3.14}
\end{aligned}$$

The derivatives of v with respect to r in formula (3.14) are computed according to the homogeneous expressions of (3.10–3.13). One advantage of decomposition (3.14) is that it reduces redundant computations. Observe that the inhomogeneous contribution contains no terms which are associated with the input functions $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1) = (v, \frac{\partial v}{\partial n}) \Big|_\Gamma$. The computational savings that result from this observation are two-fold. On the one hand, we can compute the inhomogeneous contribution once and then, for each pair of input functions $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)$, simply add it to the result of the homogeneous extension separately. On the other hand, if we desire to solve a problem with a different right-hand side f , then we do not need to redo the homogeneous extensions, but only the new inhomogeneous contribution. This latter observation implies that solving multiple related problems will be very efficient, which we explore in the numerical tests of Section 3.3.8.

In the case that $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)$ happens to be the trace of a function v which satisfies the inhomogeneous Helmholtz equation (3.1a) on Ω , i.e., $\xi_0 = v|_\Gamma$ and $\xi_1 = \frac{\partial v}{\partial n} \Big|_\Gamma$, then formula (3.7) or, equivalently, (3.14), yields a fifth-order Taylor approximation of v . We emphasize, however, that these formulae can be applied to any function $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)$ given at the boundary Γ , and we will actually apply them to the specially chosen basis functions which are not traces of any solution to the Helmholtz equation. This allows us to incorporate the boundary condition (3.1b) into the discrete BEP (3.4). To that end, we consider the extension process defined by (3.14) as an affine transformation \mathbf{Ex} , which maps a pair of functions $\xi_\Gamma = (\xi_0, \xi_1) \Big|_\Gamma$ defined on the circle Γ to a new function ξ_γ defined at the nodes of the discrete boundary γ :

$$\xi_\gamma = \mathbf{Ex}\boldsymbol{\xi}_\Gamma = \mathbf{Ex}(\xi_0, \xi_1) \Big|_\gamma = \mathbf{Ex}_H(\xi_0, \xi_1) \Big|_\gamma + \mathbf{Ex}_I f, \tag{3.15}$$

where \mathbf{Ex}_H and \mathbf{Ex}_I are the homogeneous and inhomogeneous contribution of (3.14), respectively. Observe that while \mathbf{Ex} is an affine mapping, \mathbf{Ex}_H is a linear operator with respect to its argument $\boldsymbol{\xi}_\Gamma$.

Note, that to ensure the design convergence rate of the overall method, the number of terms in the Taylor formula (3.7) that defines the operator $\mathbf{E}\mathbf{x}$ must match in a particular way the accuracy of the finite difference scheme employed inside the computational domain. The corresponding relations follow from the approximation theorems of the continuous potentials of elliptic operators by difference potentials proven by Reznik [38,39]. While the actual proofs are delicate, the end results can be used in their own right, and some of those results are reproduced in [18, Section 4.4]. Moreover, in practice it turns out that taking fewer terms in (3.7) than prescribed by [38] may be sufficient. For the fourth-order accurate scheme (2.33) that we have implemented for the current study, the original Reznik theorem would suggest taking six terms, which is the sum of the accuracy of the scheme (fourth order) and the order of the differential equation (second). However, it appears sufficient to truncate the Taylor expansion after the fourth derivative term to preserve the overall fourth-order accuracy. For other finite difference schemes, the number of terms in the expansion will need to be chosen accordingly, see [18,38,39] for more detail.

3.2.2 Series representation of boundary functions

The next step is to select a basis for the space of smooth pairs of functions, $\boldsymbol{\xi}_\Gamma$, on Γ :

$$\boldsymbol{\psi}_n^{(0)} = (\psi_n^{(0)}, 0) \quad \text{and} \quad \boldsymbol{\psi}_n^{(1)} = (0, \psi_n^{(1)}), \quad n = -\infty, \dots, \infty. \quad (3.16)$$

This basis will help us represent the data $(u, \frac{\partial u}{\partial n})|_\Gamma$ of the solution u to problem (3.1), see Section 3.2.3. Provided that the expansion of a given $\boldsymbol{\xi}_\Gamma$ with respect to system (3.16) converges sufficiently fast, we may truncate it and replace the infinite series with a finite sum:

$$\boldsymbol{\xi}_\Gamma = \underbrace{\sum_{n=-N}^N c_n^{(0)} \boldsymbol{\psi}_n^{(0)}}_{(\xi_0, 0)} + \underbrace{\sum_{n=-N}^N c_n^{(1)} \boldsymbol{\psi}_n^{(1)}}_{(0, \xi_1)}, \quad (3.17)$$

where the number N that would guarantee the desired accuracy can be taken relatively small. This is the case, e.g., for the Fourier series when $\boldsymbol{\xi}_\Gamma$ are smooth periodic functions (see Section 3.2.4). However, a different basis (3.16) may also be chosen, and the particular cases of both Fourier and Chebyshev bases are invoked and discussed in subsequent examples to suit the goal of solving specific non-standard boundary value problems (3.1). In addition, we emphasize that the basis (3.16) may be selected independently of the discretization grid N_0 , and this is accomplished by choosing the accuracy of representation (3.17) ahead of time so that it will match or exceed the accuracy that can be achieved on the grid (this is done for our case in Section 3.3.2).

In the interest of solving problem (3.1), we will first apply the operator \mathbf{Ex} of Section 3.2.1 and extend an arbitrary ξ_Γ represented by the truncated series (3.17), from the continuous boundary Γ to the grid boundary γ . Then we substitute the resulting ξ_γ into the discrete BEP (3.4). This yields a linear system in which the unknowns will be the coefficients $c_n^{(0)}$ and $c_n^{(1)}$, $n = -N, \dots, N$.

3.2.3 Applying the extension to form a linear system

Applying the extension operator \mathbf{Ex} of (3.15) to the series representation (3.17) of a pair of functions defined on the boundary, ξ_Γ , and noting that the operator is linear, we obtain:

$$\begin{aligned}\xi_\gamma &= \mathbf{Ex}\xi_\Gamma = \mathbf{Ex} \left(\sum_{n=0}^N c_n^{(0)} \psi_n^{(0)} + \sum_{n=0}^N c_n^{(1)} \psi_n^{(1)} \right) \\ &= \sum_{n=0}^N c_n^{(0)} \mathbf{Ex}_H \psi_n^{(0)} + \sum_{n=0}^N c_n^{(1)} \mathbf{Ex}_H \psi_n^{(1)} + \mathbf{Ex}_I f.\end{aligned}\tag{3.18}$$

where $\mathbf{Ex} \psi_n^{(0)} = \mathbf{Ex} (\psi_n^{(0)}, 0)$ and $\mathbf{Ex} \psi_n^{(1)} = \mathbf{Ex} (0, \psi_n^{(1)})$ according to (3.16). In the homogeneous case (i.e., where $f = 0$ and thus $\mathbf{Ex}_I f = 0$ in the above), substituting expression (3.18) into the discrete BEP (3.4) yields a system of linear algebraic equations:

$$\sum_{n=-N}^N c_n^{(0)} P_\gamma \mathbf{Ex} \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} P_\gamma \mathbf{Ex} \psi_n^{(1)} = \sum_{n=-N}^N c_n^{(0)} \mathbf{Ex} \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} \mathbf{Ex} \psi_n^{(1)},$$

which we formalize by gathering the corresponding basis terms on the left-hand side:

$$\sum_{n=-N}^N c_n^{(0)} (P_\gamma - I_\gamma) \mathbf{Ex} \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} (P_\gamma - I_\gamma) \mathbf{Ex} \psi_n^{(1)} = 0.\tag{3.19}$$

If instead we solve an inhomogeneous problem, we substitute (3.18) into the inhomogeneous discrete BEP (3.5) to yield the corresponding system, and the result of this substitution and gathering the basis terms together is as follows:

$$\sum_{n=0}^N c_n^{(0)} (P_\gamma - I_\gamma) \mathbf{Ex}_H \psi_n^{(0)} + \sum_{n=0}^N c_n^{(1)} (P_\gamma - I_\gamma) \mathbf{Ex}_H \psi_n^{(1)} = -\text{Tr} \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)} - (P_\gamma - I_\gamma) \mathbf{Ex}_I f,\tag{3.20}$$

In formulae (3.19-3.20), I_γ is the identity operator in the space of grid functions ξ_γ defined on γ . In matrix form, the linear system (3.19) can be recast as

$$\mathbf{Qc} = \mathbf{0},\tag{3.21}$$

where the matrix \mathbf{Q} is given by

$$\mathbf{Q} = \underbrace{[(\mathbf{P}_\gamma - \mathbf{I}_\gamma)\mathbf{E}\mathbf{x}\psi_{-N}^{(0)}, \dots, (\mathbf{P}_\gamma - \mathbf{I}_\gamma)\mathbf{E}\mathbf{x}\psi_N^{(0)}]}_{\mathbf{Q}_0}, \underbrace{[(\mathbf{P}_\gamma - \mathbf{I}_\gamma)\mathbf{E}\mathbf{x}\psi_{-N}^{(1)}, \dots, (\mathbf{P}_\gamma - \mathbf{I}_\gamma)\mathbf{E}\mathbf{x}\psi_N^{(1)}]}_{\mathbf{Q}_1}, \quad (3.22)$$

and \mathbf{c} is a vector of unknown coefficients, $\mathbf{c} = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$. Similarly, the matrix form of the inhomogeneous equation (3.20) is given by

$$\mathbf{Q}_H \mathbf{c} = -\text{Tr} \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)} - \mathbf{Q}_I f, \quad (3.23)$$

in which \mathbf{Q}_H is the same as \mathbf{Q} (3.22) in the homogeneous case, and the term \mathbf{Q}_I is defined by $\mathbf{Q}_I f = (\mathbf{P}_\gamma - \mathbf{I}_\gamma)\mathbf{E}\mathbf{x}_I f$.

The dimension of the matrix \mathbf{Q} in (3.22), is $|\gamma| \times 2(2N + 1)$, where $|\gamma|$ is the total number of nodes in the grid boundary γ , and the dimension of the vector \mathbf{c} is $2(2N + 1)$. The first $2N + 1$ columns of the matrix \mathbf{Q} form the sub-matrix \mathbf{Q}_0 and correspond to the coefficients $c_n^{(0)}$, $n = -N, \dots, N$, while the last $2N + 1$ columns of \mathbf{Q} form the sub-matrix \mathbf{Q}_1 and correspond to the coefficients $c_n^{(1)}$, $n = -N, \dots, N$. These dimensions are unchanged in the inhomogeneous case, with the dimensions of \mathbf{Q}_H matching those of \mathbf{Q} , and the right-hand side terms of (3.23) resolves to a single vector of length $|\gamma|$.

Any solution $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}]$ to the linear system (3.21) furnishes ξ_Γ via formula (3.17), and this, in turn, yields $\xi_\gamma = \mathbf{E}\mathbf{x}\xi_\Gamma$. As follows from the results of [38], the corresponding difference potential (3.2) with the density ξ_γ provides a fourth-order accurate approximation to the continuous Calderon potential of the Helmholtz operator \mathbf{L} with the density ξ_Γ , see also [8, Part III, Chapter 1] and [18, Section 4.4]. Moreover, the continuous Calderon potential u solves the homogeneous Helmholtz equation (3.1a) on Ω , and the density ξ_Γ of (3.17) approximates its data $(u, \frac{\partial u}{\partial n})|_\Gamma$ on Γ . This is also the case when the equation is inhomogeneous: the solution to the inhomogeneous linear system (3.23) provides the coefficients to approximate $\xi_\gamma = \mathbf{E}\mathbf{x}\xi_\Gamma$ via formula (3.17), which, by design, solves the inhomogeneous BEP (3.5), and the solution of the inhomogeneous Helmholtz equation is given by the inhomogeneous potential (3.6).

However, the linear system (3.21) may have multiple solutions. Indeed, it does not take into account the boundary condition (3.1b) because it is derived from the discrete BEP (3.4) only. To take the boundary condition into account and make sure the overall solution is unique, system (3.21) needs to be modified and/or supplemented by additional equations. These remarks, as well as the analysis that follows for the homogeneous case, apply immediately to the inhomogeneous problem also with no more modification than the inclusion of the inhomogeneous terms of (3.23). The simplest cases to analyze are those of the Dirichlet and Neumann boundary conditions, which we will now present.

Simple cases: Dirichlet and Neumann problems When the boundary condition is Dirichlet, equation (3.1b) reduces to $u|_{\Gamma} = \phi_{\Gamma}$, and we can expand the given Dirichlet data function ϕ_{Γ} (i.e., the first component of the data) with respect to the chosen basis (3.16), yielding the coefficients $c_n^{(0)}$, $n = -N, \dots, N$. Recalling that $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}] = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$ in (3.21), the vector $\mathbf{c}^{(0)}$ in the Dirichlet case can be considered given while the vector $\mathbf{c}^{(1)}$ is unknown. Accordingly, system (3.21) is recast as $\mathbf{Q}_1 \mathbf{c}^{(1)} = -\mathbf{Q}_0 \mathbf{c}^{(0)}$, where the right-hand side can now be thought of as a given vector of dimension $|\gamma|$. Then, the system is solved for $\mathbf{c}^{(1)}$ in the sense of least squares.

We may choose the number of basis functions N independent of the size of the discretization grid \mathbb{N}_0 . This number can be fixed so that the accuracy of the truncated expansion (3.17) at the boundary would exceed any accuracy that one might expect to obtain on the grid. This is easy to achieve at a moderate cost, because when the boundary data are smooth and periodic, their Fourier expansion converges rapidly (see footnote 3 on page 44) and the resulting N appears not very large. Numerical studies on how to efficiently choose the number of basis functions for each grid is considered in Section 3.3. Once the dimension N of the boundary representation (3.17) has been fixed, the final accuracy of the solution on the domain is controlled only by the size of the grid \mathbb{N}_0 . For sufficiently fine grids, one should typically expect $|\gamma| \gg 2N + 1$. Hence, the system $\mathbf{Q}_1 \mathbf{c}^{(1)} = -\mathbf{Q}_0 \mathbf{c}^{(0)}$ is overdetermined and admits a robust solution by least squares. Moreover, as the original boundary value problem (3.1) has a unique solution, the discrete least squares solution is “almost classical” in the sense that as the grid is refined, the residual at the minimum decreases to zero with the rate determined by the accuracy of the finite difference approximation.

In the Neumann case, conversely, the vector $\mathbf{c}^{(1)}$ is given by expanding the boundary condition $\frac{\partial u}{\partial n} = \phi|_{\Gamma}$ with respect to the basis (3.16), the vector $\mathbf{c}^{(0)}$ is unknown, and the system $\mathbf{Q}_0 \mathbf{c}^{(0)} = -\mathbf{Q}_1 \mathbf{c}^{(1)}$ is solved in the sense of the least squares for $\mathbf{c}^{(0)}$. In any case, once both $\mathbf{c}^{(0)}$ and $\mathbf{c}^{(1)}$ are known, the vector $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}]$ is substituted into (3.17), and the resulting ξ_{Γ} is extended to γ via the extension operator: $\xi_{\gamma} = \mathbf{E} \mathbf{x} \xi_{\Gamma}$. The difference potential (3.2) with density ξ_{γ} is then computed to approximate the solution u to the boundary value problem (3.1) with fourth-order accuracy on the grid \mathbb{N}^+ .

A more detailed analysis of the Dirichlet and Neumann boundary conditions is provided in [18] with numerical verification on both a circular and an elliptical boundary, as well as an analysis of a simple Robin boundary condition with constant coefficients (but without corresponding numerical computations). In Sections 3.2.4 and 3.2.5, we present a more comprehensive analysis of the Robin boundary condition, which includes variable and/or discontinuous coefficients. In particular, this allows us to consider mixed boundary conditions, e.g., Dirichlet on one part of the boundary and Neumann on the other part of the boundary. As stated previously, this

analysis holds for the inhomogeneous case as well with the addition of the inhomogeneous terms of (3.23). We present numerical verification of each of these BC types for both homogeneous and inhomogeneous problems in Section 3.3.

Finally, we emphasize that the reduction of problem (3.1) from its domain to the boundary based on Calderon's boundary equations with projections, regardless of the type of the boundary condition (3.1b), is always well posed as long as the original problem (3.1) itself is well posed, see [8, 18]. This is in contrast to methods based on boundary integral equations, for which care must be exercised, on a case-by-case basis, in choosing the equivalent boundary sources such that the resulting Fredholm integral equation is of the second kind (well-posed) rather than the first kind (ill-posed).

3.2.4 The Robin boundary condition with smooth variable coefficients

Consider the case that formula (3.1b) represents a general Robin boundary condition with variable coefficients:

$$\alpha(\theta)u(\theta) + \beta(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) = \phi(\theta), \quad (3.24)$$

where α , β , and ϕ are smooth periodic functions of the polar angle $\theta \in [0, 2\pi]$. We expand each term of (3.24) with respect to the chosen basis (3.16), and obtain a set of linear algebraic equations that will supplement system (3.21). In doing so, it will be convenient, though not necessary, to consider the same basis functions for both u and $\frac{\partial u}{\partial \mathbf{n}}$, so that $\psi_n^{(0)} = \psi_n^{(1)}$, $n = -N, \dots, N$, in formula (3.16).

Since formula (3.24) is comprised of smooth 2π -periodic functions, it is natural to choose a complex-exponential Fourier basis, $\psi_n^{(0)}(\theta) = \psi_n^{(1)}(\theta) = e^{in\theta}$. To express the left-hand side of (3.24) in this basis, we use a well-known convolution formula for the Fourier coefficients of a product of two functions. Let $f(\theta)$ and $g(\theta)$ be 2π -periodic, and denote by \hat{f}_n and \hat{g}_n their Fourier coefficients for the expansion with respect to the complex exponentials $e^{in\theta}$, $n = 0, \pm 1, \pm 2, \dots$. Then, it is easy to show that

$$(\widehat{fg})_n = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \hat{g}_m \hat{f}_{m-n}. \quad (3.25)$$

We now expand the boundary condition (3.24) using this result. Let $c_n^{(0)}$ and $c_n^{(1)}$ represent the Fourier coefficients of $u(\theta)$ and $\frac{\partial u}{\partial \mathbf{n}}(\theta)$, respectively, and let $\hat{\alpha}_n, \hat{\beta}_n$, and $\hat{\phi}_n$ be the coefficients of $\alpha(\theta), \beta(\theta)$, and $\phi(\theta)$, respectively. Then, according to (3.25), formula (3.24) becomes:

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\sum_{m=-\infty}^{\infty} \hat{\alpha}_m c_{m-n}^{(0)} + \sum_{m=-\infty}^{\infty} \hat{\beta}_m c_{m-n}^{(1)} \right) e^{in\theta} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \hat{\phi}_n e^{in\theta}.$$

By orthogonality of the basis functions, we obtain the following linear equation for each n :

$$\sum_{m=-\infty}^{\infty} \left(\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)} \right) = \hat{\phi}_n. \quad (3.26)$$

Moreover, since for sufficiently smooth functions their Fourier coefficients decay rapidly,³ it is sufficient to take only finitely many equations from (3.26) to supplement system (3.21). The summation range on the left-hand side of each of these equations can also be chosen finite. The specific number of equations needed will be determined based on the desired accuracy of the Fourier expansions.

Assume we are given a tolerance $\epsilon > 0$. Then, using the generic notations $f(\theta)$ and $g(\theta)$, we can find a positive integer number N [cf. formula (3.17)] such that

$$|\hat{f}_n| < \epsilon \text{ and } |\hat{g}_n| < \epsilon \text{ for } |n| \geq N. \quad (3.27)$$

Since the coefficients \hat{f}_n and \hat{g}_n decay rapidly, the number N in formula (3.27) is typically not large even if ϵ is taken on the order of the machine precision (specific choices that we have made for particular examples are discussed in Section 3.3.) In general, assuming that ϵ in (3.27) is small, we can replace all of the coefficients with indices $|n| \geq N$ in the Fourier expansions of f and g by zeros.

We therefore consider a pair of sufficiently smooth 2π -periodic functions $f(\theta)$ and $g(\theta)$ for which we set $\hat{f}_n = 0$ and $\hat{g}_n = 0$ for $|n| \geq N$. To find the coefficients $(\widehat{fg})_n$ of the truncated Fourier expansion for their product fg , we identify and exclude from the last sum on the right-hand side of formula (3.25) all terms for which either $\hat{f}_{m-n} = 0$ or $\hat{g}_m = 0$ (i.e., those terms, for which either $|m-n| > N$ or $|m| > N$, respectively.)

If $n \geq 0$, then $m \geq m-n$; thus, the upper bound for the summation will be $m = N$, and the lower bound will be achieved when $m-n = -N$, which, solved for m , yields $m = n-N$. Similarly, when $n < 0$, we have $m < m-n$, which results in the lower bound being reached by $m = -N$, and the upper bound is reached by $m-n = N$, which implies $m = n+N$. Hence, we have:

$$(\widehat{fg})_n = \begin{cases} \sum_{m=n-N}^N \hat{g}_m \hat{f}_{m-n}, & 0 \leq n \leq 2N, \\ \sum_{m=-N}^{n+N} \hat{g}_m \hat{f}_{m-n}, & -2N \leq n < 0. \end{cases} \quad (3.28a)$$

If $n > 2N$, then the summation range in the first sum on the right-hand side of (3.28a) becomes empty; if $n < -2N$, then the summation range becomes empty in the second sum.

³For an r -differentiable function with the derivative of order r in L_2 , the rate of decay of its Fourier coefficients is $o(n^{-r})$, and accordingly, the rate of convergence of its Fourier series is $o(n^{-(r-\frac{1}{2})})$, see, e.g., [40, Section 3.1.3].

Empty summation ranges yield zero Fourier coefficients so that

$$(\widehat{fg})_n = 0, \quad |n| > 2N. \quad (3.28b)$$

Applying formulae (3.28) to equations (3.26), we arrive at the following finite system:

$$\begin{cases} \sum_{m=n-N}^N \left(\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)} \right) = \hat{\phi}_n, & 0 \leq n \leq 2N, \\ \sum_{m=-N}^{n+N} \left(\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)} \right) = \hat{\phi}_n, & -2N \leq n < 0, \end{cases} \quad (3.29)$$

where the terms $\hat{\phi}_n$ on the right-hand side become zero whenever $|n| \geq N$.

System (3.29) provides $4N + 1$ additional equations to supplement the $|\gamma|$ equations of system (3.21). The purpose of equations (3.29) is to take into account the boundary condition (3.24), which is a particular form of (3.1b), whereas system (3.21) is responsible for satisfying the differential equation (3.1a). Combining them yields a non-trivial solution which will satisfy both the differential equation and boundary condition. The only inhomogeneity of the overall system of equations comes from the coefficients $\hat{\phi}_n$ that are non-zero, i.e., from equations (3.29) with $n = -N, \dots, N$. The dimension of system (3.21), (3.29) is $[|\gamma| + (4N + 1)] \times 2(2N + 1)$, and its solution $\mathbf{c} = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$ is to be sought for in the sense of least squares.

In the previously analyzed cases of the Dirichlet and Neumann boundary conditions, this system simplifies to coincide with our prior analysis. The Dirichlet boundary condition is equivalent to $\alpha = 1$ and $\beta = 0$ in the general equation (3.24), yielding $u(\theta) = \phi(\theta)$. Then, $\hat{\alpha}_0 = 1$, $\hat{\alpha}_n = 0$ for $n = -N, \dots, -1, 1, \dots, N$, and $\hat{\beta}_n = 0$ for $n = -N, \dots, N$, so that system (3.29) reduces to

$$c_n^{(0)} = \hat{\phi}_n, \quad n = -2N, \dots, 2N.$$

Moreover, as $\hat{\phi}_n = 0$ for $|n| > N$, we can simply disregard the corresponding coefficients $c_n^{(0)}$, and keep only those $c_n^{(0)}$, for which $n = -N, \dots, N$. Substituting these $c_n^{(0)} = \hat{\phi}_n$ into (3.21), we get:

$$\mathbf{Q}_1 \begin{bmatrix} c_{-N}^{(1)} \\ \vdots \\ c_N^{(1)} \end{bmatrix} = -\mathbf{Q}_0 \begin{bmatrix} \hat{\phi}_{-N} \\ \vdots \\ \hat{\phi}_N \end{bmatrix},$$

which is to be solved by least squares with respect to the unknown $c_n^{(1)}$, $n = -N, \dots, N$.

Similarly, the Neumann boundary condition corresponds to $\alpha = 0$ and $\beta = 1$ in formula (3.24). Consequently, $\hat{\alpha}_n = 0$ for $n = -N, \dots, N$, $\hat{\beta}_1 = 1$, and $\hat{\beta}_n = 0$ for $n =$

$-N, \dots, -1, 1, \dots, N$. This reduces equations (3.29) to

$$c_n^{(1)} = \hat{\phi}_n, \quad n = -2N, \dots, 2N.$$

Again, we disregard those $c_n^{(1)}$ for which $n = -2N, \dots, -N-1, N+1, \dots, 2N$, because the corresponding $\hat{\phi}_n = 0$ for $|n| > N$. Substituting the remaining $c_n^{(1)} = \hat{\phi}_n$, $n = -N, \dots, N$, into (3.21) yields

$$\mathbf{Q}_0 \begin{bmatrix} c_{-N}^{(0)} \\ \vdots \\ c_N^{(0)} \end{bmatrix} = -\mathbf{Q}_1 \begin{bmatrix} \hat{\phi}_{-N} \\ \vdots \\ \hat{\phi}_N \end{bmatrix}.$$

This is precisely the approach explained in Section 3.2.3 for Dirichlet and Neumann BCs, and the analysis of the inhomogeneous case is equivalent with the addition of the inhomogeneous terms on the right-hand side.

In the cases other than those of the Dirichlet or Neumann boundary conditions, the full system (3.29) has $4N+1$ equations, and it is not immediately obvious whether the $2N$ homogeneous equations can likewise be disregarded. Indeed, while it is possible that for $-2N \leq n \leq -N-1$ or for $N+1 \leq n \leq 2N$ the products of small terms on the left-hand side of the corresponding equations (3.29) will be $o(\epsilon)$ [i.e., asymptotically smaller than $\mathcal{O}(\epsilon)$], it is not automatically guaranteed. Experimentally, the setting with no homogeneous equations in (3.29) was tested, and we observed that keeping or dropping those equations made very little difference for the cases computed in Section 3.3.3. Theoretically, however, this issue requires more analysis, which we leave for future study.

3.2.5 The Robin boundary problem with discontinuous variable coefficients

Consider the general Robin boundary condition (3.24), but with the relaxed assumptions that α , β , and ϕ are bounded and piecewise smooth rather than globally smooth. For simplicity, assume that $\Gamma = \Gamma_1 \cup \Gamma_2$, where Γ_1 includes all points on the circle with $\theta \in [0, a)$ and Γ_2 includes all points on the circle with $\theta \in [a, 2\pi)$ for some $0 < a < 2\pi$. Assume that α , β , and ϕ are smooth and bounded on either Γ_1 or Γ_2 , but are not necessarily continuous on the entire circle Γ . For example, they may have a jump discontinuity at $\theta = 0$ and/or $\theta = a$.

Since the boundary data are no longer assumed to be periodic, we introduce a Chebyshev basis. Consider the standard Chebyshev polynomial basis, $\{T_n(x)\}_{n=0}^{\infty}$, $x \in [-1, 1]$, with the weight $\omega(x) = 2/\pi\sqrt{1-x^2}$. For a given function $f(x)$, denote its Chebyshev coefficients by \hat{f}_n , $n = 0, 1, 2, \dots$. To express the left-hand side of the boundary condition (3.24) in the Chebyshev basis, we need to find the form of the expansion for the products $\alpha(\theta)u(\theta)$ and $\beta(\theta)\frac{\partial u}{\partial n}(\theta)$. In appendix A, the Chebyshev coefficients of a product are derived for a pair of arbitrary smooth

functions f and g on $[-1, 1]$, see formulae (A.2). To implement this expansion in practice, one first needs to truncate it and replace the series by a finite sum. As Chebyshev coefficients of smooth functions decay rapidly (the relation between the rate of decay and smoothness is similar to that for the Fourier coefficients, see footnote 3 on page 44), for a given $\epsilon > 0$ we can choose a relatively small number N such that

$$\begin{aligned} |\hat{f}_n| &< \epsilon, \quad n > N, \\ |\hat{g}_n| &< \epsilon, \quad n > N. \end{aligned}$$

We therefore, set $\hat{f}_n = \hat{g}_n = 0$ for $n > N$. Then formula (A.2a) for $n = 0$ immediately yields

$$(\widehat{fg})_0 = 2\hat{g}_0\hat{f}_0 + \sum_{m=1}^N \hat{g}_m\hat{f}_m. \quad (3.30a)$$

Consider the case $0 < n \leq N$. On the right-hand side of the last equality in (A.2b), we first replace the upper limit in the last sum by N because the factor \hat{g}_m in the product under the sum will be set to zero beyond $m = N$. Next, since $n > 0$, we notice that the index $m + n$ is the largest, and the corresponding terms become zero when $m > N - n$, which yields:

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m(\hat{f}_{n-m} + \hat{f}_{n+m}) + \hat{g}_n(\hat{f}_{2n} + 2\hat{f}_0) + \sum_{m=n+1}^N \hat{g}_m\hat{f}_{m-n} + \sum_{m=n+1}^{N-n} \hat{g}_m\hat{f}_{m+n} \right). \quad (3.30b)$$

We use equation (3.30b) when $1 \leq n < \frac{N}{2}$. Clearly the term \hat{f}_{2n} in (3.30b) is zero whenever $n > \frac{N}{2}$ because its index will be $2n > N$. Moreover, the last sum will be zero for $n \geq \frac{N}{2}$ since the smallest index of \hat{f}_{m+n} will be $m + n = \frac{N}{2} + 1 + \frac{N}{2} > N$. To write the resulting formula in a convenient manner while eliminating the zero terms for $n > \frac{N}{2}$, we first rearrange (3.30b) so that the terms with the largest indices appear last:

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m\hat{f}_{n-m} + 2\hat{g}_n\hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m\hat{f}_{m-n} + \sum_{m=0}^{n-1} \hat{g}_m\hat{f}_{m+n} + \hat{g}_n\hat{f}_{2n} + \sum_{m=n+1}^{N-n} \hat{g}_m\hat{f}_{m+n} \right).$$

Next, we eliminate terms as the index becomes larger since they become zero. If $n \geq \frac{N}{2}$ (note that $n = \frac{N}{2}$ can occur only if N is even), then the summation range in the last sum becomes empty since

$$N - n \leq N - \frac{N}{2} = \frac{N}{2} \leq n < n + 1.$$

Thus, if N is even, then we can write for $n = \frac{N}{2}$:

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{n-m} + 2\hat{g}_n \hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m \hat{f}_{m-n} + \sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{m+n} + \hat{g}_n \hat{f}_{2n} \right). \quad (3.30c)$$

Notice that the index $m+n$ in the last sum of (3.30c) exceeds N if $m > N-n$. At the same time, we see that if $\frac{N}{2} < n \leq N$, then $N-n < n-1$. Thus, replacing the upper limit in the last sum of (3.30c) by the tighter bound $N-n$, we have, for $\frac{N}{2} < n \leq N$,

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{n-m} + 2\hat{g}_n \hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m \hat{f}_{m-n} + \sum_{m=0}^{N-n} \hat{g}_m \hat{f}_{m+n} \right). \quad (3.30d)$$

For $N < n \leq 2N$, the summation range in the second to last sum in (3.30d) is empty. In addition, $\hat{g}_n = 0$ and $\hat{f}_{m+n} = 0$ for any $m \geq 0$. Consequently, only the first sum will remain, with the upper bound replaced by N since now $n-1 \geq N$:

$$(\widehat{fg})_n = \frac{1}{2} \sum_{m=1}^N \hat{g}_m \hat{f}_{n-m}, \quad n = N+1, \dots, 2N. \quad (3.30e)$$

Finally, observe for $n > 2N$ that $n-m > 2N-m > N$, which leaves no non-zero terms so that

$$(\widehat{fg})_n = 0, \quad n > 2N. \quad (3.30f)$$

Altogether, the coefficients of the truncated Chebyshev expansion for the product fg are given by equation (3.30a) for $n = 0$, (3.30b) for $1 \leq n < \frac{N}{2}$, (3.30c) for $n = \frac{N}{2}$ (note that this occurs only if N is even), (3.30d) for $\frac{N}{2} < n \leq N$, (3.30e) for $N < n \leq 2N$, and (3.30f) for $n > 2N$.

We now derive the supplementary linear system by applying formulae (3.30) to the general boundary condition (3.24). Recall that we are considering a continuous boundary Γ partitioned into two pieces, $\Gamma = \Gamma_1 \cup \Gamma_2$, which are two arcs of the circle of radius 1 on the intervals $\theta \in [0, a)$ and $\theta \in [a, 2\pi)$, respectively. We now recast the boundary condition (3.24) as

$$\begin{aligned} \alpha^{(1)}(\theta)u(\theta) + \beta^{(1)}(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) &= \phi^{(1)}(\theta) \quad \text{on } \Gamma_1, \\ \alpha^{(2)}(\theta)u(\theta) + \beta^{(2)}(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) &= \phi^{(2)}(\theta) \quad \text{on } \Gamma_2. \end{aligned} \quad (3.31)$$

To utilize the Chebyshev basis for (3.31), we perform a linear change of variables on Γ_1 and Γ_2 from θ to x so that $x \in [-1, 1)$ in each respective case. For Γ_1 , we have $\theta \in [0, a)$ and consequently,

$$x = \frac{\theta}{a} + \frac{\theta - a}{a}, \quad (3.32a)$$

whereas for Γ_2 we have $\theta \in [a, 2\pi)$, and the transformation is given by

$$x = \frac{\theta - a}{2\pi - a} + \frac{\theta - 2\pi}{2\pi - a}. \quad (3.32b)$$

We consider two independent Chebyshev bases, one on each of the arcs, Γ_1 and Γ_2 , and denote their respective dimensions by N_1 and N_2 . As such, we have two double-sets of coefficients:

$$\begin{aligned} [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}]^T &= [c_0^{(0,1)}, \dots, c_{N_1}^{(0,1)}, c_1^{(1,1)}, \dots, c_{N_1}^{(1,1)}]^T, \\ [\mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T &= [c_0^{(0,2)}, \dots, c_{N_2}^{(0,2)}, c_1^{(1,2)}, \dots, c_{N_2}^{(1,2)}]^T, \end{aligned}$$

so that $\mathbf{c} = [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}, \mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T$. Accordingly, instead of formula (3.17) we now have:

$$\xi_\Gamma = \underbrace{\sum_{n=0}^{N_1} c_n^{(0,1)} \psi_n^{(0,1)} + \sum_{n=0}^{N_2} c_n^{(0,2)} \psi_n^{(0,2)}}_{(\xi_0, 0)} + \underbrace{\sum_{n=0}^{N_1} c_n^{(1,1)} \psi_n^{(1,1)} + \sum_{n=0}^{N_2} c_n^{(1,2)} \psi_n^{(1,2)}}_{(0, \xi_1)},$$

where

$$\begin{aligned} \psi_n^{(0,1)} &= \begin{cases} (T_n, 0) & \text{on } \Gamma_1, \\ (0, 0) & \text{on } \Gamma_2, \end{cases} & \psi_n^{(0,2)} &= \begin{cases} (0, 0) & \text{on } \Gamma_1, \\ (T_n, 0) & \text{on } \Gamma_2, \end{cases} \\ \psi_n^{(1,1)} &= \begin{cases} (0, T_n) & \text{on } \Gamma_1, \\ (0, 0) & \text{on } \Gamma_2, \end{cases} & \psi_n^{(1,2)} &= \begin{cases} (0, 0) & \text{on } \Gamma_1, \\ (0, T_n) & \text{on } \Gamma_2. \end{cases} \end{aligned} \quad (3.33)$$

The extension of a given basis function from (3.33) to the discrete boundary γ is done according to the same formulae derived for the Taylor expansion in Section 3.2.1. The matrix \mathbf{Q} is now partitioned into 4 blocks rather than 2, and will have the dimension $|\gamma| \times [2(N_1 + 1) + 2(N_2 + 1)]$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_0^{(1)} & \mathbf{Q}_1^{(1)} & \mathbf{Q}_0^{(2)} & \mathbf{Q}_1^{(2)} \end{bmatrix}. \quad (3.34)$$

As in the previous case (Section 3.2.4), the corresponding homogeneous linear system (3.21) with the matrix \mathbf{Q} of (3.34) accounts for the differential equation (3.1a) but not for the boundary condition (3.1b). Thus, we supplement it with additional equations to account for the boundary conditions.

Transforming the boundary condition (3.31) according to (A.1) and taking into account formulae (3.30) for the coefficients of the truncated Chebyshev expansion of a product of two functions, we obtain a set of additional linear equations for each Γ_i , $i = 1, 2$. Specifically, we

have for $n = 0$:

$$c_0^{(0,i)} \hat{\alpha}_0^{(i)} + \frac{1}{2} \sum_{m=1}^{N_i} c_m^{(0,i)} \hat{\alpha}_m^{(i)} + c_0^{(1,i)} \hat{\beta}_0^{(i)} + \frac{1}{2} \sum_{m=1}^{N_i} c_m^{(1,i)} \hat{\beta}_m^{(i)} = \hat{\phi}_0^{(i)}, \quad (3.35a)$$

for $1 \leq n < \frac{N_i}{2}$:

$$\begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \left(\hat{\alpha}_{n-m}^{(i)} + \hat{\alpha}_{n+m}^{(i)} \right) + c_n^{(0,i)} \left(\hat{\alpha}_{2n}^{(i)} + 2\hat{\alpha}_0^{(i)} \right) \right. \\ & + \sum_{m=n+1}^{N_i} c_m^{(0,i)} \hat{\alpha}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(0,i)} \hat{\alpha}_{m+n}^{(i)} \\ & + \sum_{m=0}^{n-1} c_m^{(1,i)} \left(\hat{\beta}_{n-m}^{(i)} + \hat{\beta}_{n+m}^{(i)} \right) + c_n^{(1,i)} \left(\hat{\beta}_{2n}^{(i)} + 2\hat{\beta}_0^{(i)} \right) \\ & \left. + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(1,i)} \hat{\beta}_{m+n}^{(i)} \right) = \hat{\phi}_n^{(i)}, \end{aligned} \quad (3.35b)$$

for $n = \frac{N_i}{2}$ (if N_i is even):

$$\begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \left(\hat{\alpha}_{n-m}^{(i)} + \hat{\alpha}_{n+m}^{(i)} \right) + c_n^{(0,i)} \left(\hat{\alpha}_{2n}^{(i)} + 2\hat{\alpha}_0^{(i)} \right) + \sum_{m=n+1}^{N_i} c_m^{(0,i)} \hat{\alpha}_{m-n}^{(i)} \right. \\ & \left. + \sum_{m=0}^{n-1} c_m^{(1,i)} \left(\hat{\beta}_{n-m}^{(i)} + \hat{\beta}_{n+m}^{(i)} \right) + c_n^{(1,i)} \left(\hat{\beta}_{2n}^{(i)} + 2\hat{\beta}_0^{(i)} \right) + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} \right) = \hat{\phi}_n^{(i)}, \end{aligned} \quad (3.35c)$$

for $\frac{N_i}{2} < n \leq N_i$:

$$\begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \hat{\alpha}_{n-m}^{(i)} + 2c_n^{(0,i)} \hat{\alpha}_0^{(i)} + \sum_{m=n+1}^{N_i} c_m^{(0,1)} \hat{\alpha}_{m-n}^{(1)} + \sum_{m=n+1}^{N_i-n} c_m^{(0,i)} \hat{\alpha}_{m+n}^{(i)} \right. \\ & \left. + \sum_{m=0}^{n-1} c_m^{(1,i)} \hat{\beta}_{n-m}^{(i)} + 2c_n^{(1,i)} \hat{\beta}_0^{(i)} + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(1,i)} \hat{\beta}_{m+n}^{(i)} \right) = \hat{\phi}_n^{(i)}, \end{aligned} \quad (3.35d)$$

and for $N_i < n \leq 2N_i$:

$$\frac{1}{2} \left(\sum_{m=0}^{N_i} c_m^{(0,i)} \alpha_{n-m}^{(i)} + \sum_{m=0}^{N_i} c_m^{(1,i)} \beta_{n-m}^{(i)} \right) = 0. \quad (3.35e)$$

This gives us a total of $2N_i + 1$ extra equations for the $2(N_i + 1)$ coefficients $[\mathbf{c}^{(0,i)}, \mathbf{c}^{(1,i)}]$, where $i \in \{1, 2\}$, so that the augmented system comprised of (3.21) with \mathbf{Q} given by (3.34) and

equations (3.35) will have the dimension $(|\gamma| + 2N_1 + 1 + 2N_2 + 1) \times [2(N_1 + 1) + 2(N_2 + 1)]$. It shall be solved in the sense of least squares. We also observe that the only inhomogeneity in the overall system arises from the supplemental equations (3.35a–3.35d) which account for the boundary condition, whereas the remaining supplemental equations (3.35e) are homogeneous.

Finally, we demonstrate that in the case of simple boundary conditions, such as Dirichlet or Neumann, the supplemental equations reduce to a diagonal system. For example, consider a Dirichlet boundary condition on Γ_1 and a Neumann boundary condition on Γ_2 . Then,

$$\begin{aligned}\alpha^{(1)}(\theta) &= 1, \beta^{(1)}(\theta) = 0, \\ \alpha^{(2)}(\theta) &= 0, \beta^{(2)}(\theta) = 1.\end{aligned}$$

Thus the Chebyshev coefficients of $\alpha^{(i)}$ and $\beta^{(i)}$, $i = 1, 2$, will be:

$$\begin{aligned}\hat{\alpha}_0^{(1)} &= 1, \hat{\alpha}_n^{(1)} = 0 \text{ for } n = 1, \dots, N, \\ \hat{\beta}_n^{(1)} &= 0 \text{ for } n = 0, \dots, N, \\ \hat{\alpha}_n^{(2)} &= 0 \text{ for } n = 0, \dots, N, \\ \hat{\beta}_0^{(2)} &= 1, \hat{\beta}_n^{(2)} = 0 \text{ for } n = 1, \dots, N.\end{aligned}$$

Consequently, equations (3.35) for $i = 1$ reduce to

$$\begin{aligned}c_0^{(0,1)} &= \hat{\phi}_0^{(1)} \text{ for } n = 0 \text{ from (3.35a),} \\ c_n^{(0,1)} &= \hat{\phi}_n^{(1)} \text{ for } n = 1, \dots, N_1 \text{ from (3.35b)–3.35d,} \\ 0 &= 0 \text{ for } n = N_1 + 1, \dots, 2N_1 \text{ from (3.35e),}\end{aligned}\tag{3.36}$$

while for $i = 2$ they reduce to

$$\begin{aligned}c_0^{(1,2)} &= \hat{\phi}_0^{(2)} \text{ for } n = 0 \text{ from (3.35a),} \\ c_n^{(1,2)} &= \hat{\phi}_n^{(2)} \text{ for } n = 1, \dots, N_1 \text{ from (3.35b)–3.35d,} \\ 0 &= 0 \text{ for } n = N_1 + 1, \dots, 2N_1 \text{ from (3.35e).}\end{aligned}\tag{3.37}$$

Subsequently, we substitute (3.36) and (3.37) into system (3.21) as we did in the Fourier case (Section 3.2.4) to obtain a reduced system for the coefficients that remain unknown, $\mathbf{c}^{(1,1)}$ and $\mathbf{c}^{(0,2)}$ that is also solved in the sense of least squares. Clearly, any combination of the Dirichlet and Neumann boundary conditions on Γ_1 and Γ_2 will yield a similar reduced system, and also any inhomogeneity of the equation (3.1a) will require only the addition of the inhomogeneous terms of the discrete BEP (3.23).

To conclude this section we mention that the same considerations as outlined in Section 3.2.4 for the Fourier case, apply to the Chebyshev case as well. Namely, the dimensions N_1 and N_2

should be chosen so as to have the accuracy of the truncated Chebyshev expansion on either Γ_1 or Γ_2 exceed the accuracy attainable on the grid, and the specific choices that we have made are discussed in Section 3.3. Moreover, other than for the simple Dirichlet and Neumann cases it is not clear whether the homogeneous supplemental equations (3.35e) can be dropped from the overall system — this question requires further theoretical inquiry.

Another avenue for further consideration is the potential for the undesirable growth in the total number of basis functions needed when splitting the boundary Γ into segments, and using an independent basis for each piece. At first glance, it appears that we would accumulate basis functions linearly with respect to the number of partitions, eventually sacrificing the efficiency of the method. In fact, this should not be the case. A priori, any function on the continuous boundary Γ will experience less variation on a subinterval $\Gamma_i \subset \Gamma$ than on the whole of Γ , even if the boundary data are oscillatory. Thus, achieving the same level of accuracy by the expansion will require fewer coefficients on Γ_i than it would on Γ , and this property would reduce the accumulation of basis functions due to the splitting of Γ into smaller subintervals. Specific quantitative estimates along these lines will be a subject for future study.

3.2.6 Derivatives of the Chebyshev polynomials near the endpoints

The extension from the continuous boundary Γ to the discrete boundary γ via the Taylor formula (3.7) is done independently for each basis function (see Section 3.2.2), and this process requires us to provide the basis functions themselves as well as their tangential derivatives up to the fourth order, see formulae (3.12) and (3.13).⁴ For the Chebyshev basis it is well-known that the derivatives of the polynomials $T_n(x)$ near the endpoints $x = \pm 1$ are not singular, but their values become large. For example, the first derivative of the n^{th} polynomial is

$$T'_n(x) = (\cos(n \arccos x))' = \frac{n \sin(n \arccos x)}{\sqrt{1-x^2}},$$

and taking successive derivatives will clearly retain the term $1-x^2$ in the denominator with increasingly higher exponents. Inevitably, whether by chance or by sufficient refinement of the grid, we will need to compute the values of these derivatives “close” to the endpoints $x = 1$ and/or $x = -1$. Specifically, this happens when the foot of the normal dropped from a given node of γ to Γ , see Section 3.2.1, appears to be close to one of the points that partition Γ into segments ($\theta = 0$ or $\theta = a$ in Section 3.2.5). In this case, the overall accuracy may deteriorate via the loss of significant digits. We have, in fact, computationally observed such a loss of accuracy.

To avoid this undesirable phenomenon, we employ an approach that allows us to completely eliminate the need to compute the derivatives of the Chebyshev basis functions near the end-

⁴Additionally, for a higher-order scheme one may need to use a higher-order Taylor formula, requiring even higher-degree tangential derivatives to be supplied.

points. The key idea of the approach is to use an extended interval for the Chebyshev basis. That is, instead of linearly transforming the intervals $\theta \in [0, a)$ and $\theta \in [a, 2\pi)$, i.e., the arcs of the circle Γ_1 and Γ_2 , to the interval $x \in [-1, 1)$ to form the Chebyshev expansion, see formulae (3.32), we will instead linearly transform them to a smaller interval $x \in [-1 + \varepsilon, 1 - \varepsilon)$ where $\varepsilon > 0$:

$$x = \left(\frac{\theta}{a} + \frac{\theta - a}{a} \right) (1 - \varepsilon) \quad (3.38a)$$

and

$$x = \left(\frac{\theta - a}{2\pi - a} + \frac{\theta - 2\pi}{2\pi - a} \right) (1 - \varepsilon). \quad (3.38b)$$

In doing so, formulae (3.38a) and (3.38b) obviously provide a transformation between the full interval $x \in [-1, 1)$ and the two extended intervals of the variable θ :

$$x \in [-1, 1) \longleftrightarrow \theta \in [-a\sigma, a + a\sigma) \quad (3.39a)$$

and

$$x \in [-1, 1) \longleftrightarrow \theta \in [a - (2\pi - a)\sigma, 2\pi + (2\pi - a)\sigma), \quad (3.39b)$$

respectively, where $\sigma = \frac{1}{2} \frac{\varepsilon}{1 - \varepsilon} > 0$.

We then extend all the functions that define the problem, $\alpha^{(1)}(\theta)$, $\beta^{(1)}(\theta)$, $\phi^{(1)}(\theta)$ and $\alpha^{(2)}(\theta)$, $\beta^{(2)}(\theta)$, $\phi^{(2)}(\theta)$, see formula (3.31), smoothly but otherwise arbitrarily from their respective intervals $[0, a)$ and $[a, 2\pi)$ to the extended intervals (3.39a) and (3.39b), so that they can subsequently be represented as functions of x using a standard Chebyshev series on $[-1, 1)$. We also formally assume that the unknown functions u and $\frac{\partial u}{\partial n}$ are defined on the same extended intervals (3.39) so that we can identically reproduce all the arguments of Section 3.2.5 and obtain the same supplemental equations (3.35). At the same time, to build the extension (3.7), (3.8) of a given basis function from Γ to γ we will need to know this function only on the corresponding arc $\Gamma_1 \Leftrightarrow \theta \in [0, a)$ or $\Gamma_2 \Leftrightarrow \theta \in [a, 2\pi)$, or, equivalently, on the reduced interval $[-1 + \varepsilon, 1 - \varepsilon)$, because it is those actual arcs where the normals from γ can meet Γ , whereas the “tails” $x \in [-1, -1 + \varepsilon)$ and $x \in [1 - \varepsilon, 1)$ are artificial. Hence, the extension operator ***E*** x will never require any information from these tails and there will be no need to compute the derivatives of the basis functions near the endpoints (i.e., closer than ε to endpoints). It is only the reduced interval $[-1 + \varepsilon, 1 - \varepsilon)$ that will eventually contribute to system (3.21) which represents the discrete BEP (3.4).

Let $f(x)$ be a smooth function defined on $[-1 + \varepsilon, 1 - \varepsilon]$, where $0 < \varepsilon < 1$, and let $\tilde{f}(x)$ be any smooth extension of $f(x)$ to the full interval $[-1, 1]$ (i.e., $\tilde{f}(x) \equiv f(x)$ for $|x| \leq 1 - \varepsilon$, and $\tilde{f}(x)$ is smooth on all of $[-1, 1]$). Then, we expand \tilde{f} in the Chebyshev basis, see (A.1):

$$\tilde{f}(x) = \sum_{n=0}^{\infty} \hat{f}_n T_n(x), \quad x \in [-1, 1], \quad (3.40)$$

Since series (3.40) converges uniformly, then, clearly, the same convergence takes place on any subinterval, which implies, in particular, that

$$f(x) = \sum_{n=0}^{\infty} \hat{f}_n T_n(x), \quad x \in [-1 + \epsilon, 1 - \epsilon]. \quad (3.41)$$

We emphasize that even though the coefficients \hat{f}_n of the series (3.40) depend on what particular extension of $f(x)$ from $[-1 + \epsilon, 1 - \epsilon]$ to $[-1, 1]$ we choose, this dependence manifests itself only through the fact that the sum of the series (3.40) will vary on the tails $1 - \epsilon < |x| \leq 1$. At the same time, on the central sub-interval $|x| \leq 1 - \epsilon$ the sum of the series (3.41) remains the same, i.e., equal to the original $f(x)$, regardless of the specific behavior of $\tilde{f}(x)$ for $1 - \epsilon < |x| \leq 1$.

Accordingly, the Chebyshev coefficients $\hat{\alpha}_n^{(i)}$, $\hat{\beta}_n^{(i)}$, and $\hat{\phi}_n^{(i)}$, where $i = 1, 2$, will depend on the respective extensions of the functions $\alpha^{(i)}$, $\beta^{(i)}$, and $\phi^{(i)}$ from $[-1 + \epsilon, 1 - \epsilon]$ to $[-1, 1]$. As these coefficients provide the data for the supplemental equations (3.35), the solution $\mathbf{c} = [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}, \mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T$ of the overall system (3.21), (3.35), where the matrix \mathbf{Q} is given by (3.34), will also be affected by what extension of $\alpha^{(i)}$, $\beta^{(i)}$, and $\phi^{(i)}$, $i = 1, 2$, has been chosen. However, the resulting variation of \mathbf{c} will, again, correspond only to the variation of u and $\frac{\partial u}{\partial n}$ on the extension tails $1 - \epsilon < |x| \leq 1$ of both arcs, Γ_1 and Γ_2 , and will not affect the solution on the interior subinterval $|x| \leq 1 - \epsilon$, i.e., on the actual arcs themselves.

In other words, what we essentially do is enforce the boundary conditions (3.31) on the extended intervals (3.39), while the differential equation (3.1a) is still enforced on the original boundary $\Gamma = \Gamma_1 \cup \Gamma_2$ through the discrete BEP (3.4). In doing so, extension (3.39) keeps the original boundary data, and hence the definition of the problem, unaffected. At the same time, the redundancy that we build into the treatment of the boundary conditions by using the extensions allows us to circumvent the difficulties in computing the derivatives of the Chebyshev basis functions near the endpoints. The practical choice of the tolerance ϵ is discussed in Section 3.3.2.

The only remaining question is how to actually obtain the extended function \tilde{f} for a given f . Recall that the behavior of the tails of \tilde{f} (i.e., the part for which $|x| > 1 - \epsilon$) will not affect the convergence of the series (3.41) to $f(x)$ for $x \in [-1 + \epsilon, 1 - \epsilon]$. Thus we are free to choose any smooth, bounded extension of f . In the case that we are given an analytic formula for $f(x)$, $x \in [-1 + \epsilon, 1 - \epsilon]$, which also defines a smooth function on $[-1, 1]$, we can simply use the same formula on the larger interval. Otherwise, we can employ a polynomial extension of order J :

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [-1 + \epsilon, 1 - \epsilon], \\ \sum_{j=0}^J \frac{1}{j!} \frac{d^j f(-1+\epsilon)}{dx^j} (x + 1 - \epsilon)^j, & x < -1 + \epsilon, \\ \sum_{j=0}^J \frac{1}{j!} \frac{d^j f(1-\epsilon)}{dx^j} (x - 1 + \epsilon)^j, & x > 1 - \epsilon. \end{cases} \quad (3.42)$$

Formula (3.42) guarantees that the extended function \tilde{f} is smooth and hence bounded on $[-1, 1]$.

In Sections 3.3.5 and 3.3.9, we compare the performance of the algorithm with and without use of the extended Chebyshev intervals. Our computations convincingly corroborate that the proposed approach completely eliminates the adverse numerical effect of having large derivatives near the endpoints. Therefore, we did not feel necessary to look for any alternatives. Yet one may, of course, use other strategies as well. For example, another well known orthogonal system that guarantees rapid convergence of the expansion for smooth non-periodic functions is the Legendre polynomials [41, Appendix B.1]. The behavior of the derivatives of the original Legendre polynomials near the endpoints is similar to that of the Chebyshev polynomials. However, the associated Legendre functions vanish near the endpoints along with a certain number of derivatives [42, Section 18.11]. hence, they can potentially be used, although this approach will require a further inquiry. Yet one more alternative approach may be to use a mapping which alleviates the density of the Chebyshev nodes near the endpoints, see [43, 44].

3.2.7 Structure of the algorithm

From the discussion in Sections 3.2.1 through 3.2.5 we see that the entire computational procedure can fundamentally be split into two parts. The first part involves selecting the basis on the boundary Γ , see (3.16) or (3.33); extending the individual basis functions from Γ to γ with the help of the homogeneous contribution of the extension operator $\mathbf{E}x_H$, see (3.7), (3.8); applying the discrete projection (3.3) by solving the AP (2.52); and eventually obtaining the linear system (3.21) based on the discrete BEP (3.4) [the matrix \mathbf{Q} is given by (3.22) or (3.34)]. This part does not involve the boundary condition (3.1b) in any way, and hence does not change when this boundary condition changes. In particular, one and the same system (3.21) will work for any boundary condition of type (3.24), whether it be pure Dirichlet, pure Neumann, or general Robin boundary condition with smooth α , β , and ϕ . The only differences for an inhomogeneous problem are that inhomogeneous contribution to the extension $\mathbf{E}x_I$ must be applied to the right-hand side f of (3.1a); the inhomogeneous projection then included in the formation of the linear system (3.23), which is based on the inhomogeneous

The second part of the algorithm accounts for the boundary conditions via the supplemental equations (3.29) or (3.35). It is this part only that changes when the boundary condition changes, while otherwise the algorithm stays intact. Of course, if the boundary condition involves a discontinuity, see Section 3.2.5, and the boundary is accordingly partitioned into segments, then

the matrix \mathbf{Q} of (3.34) also needs to be recomputed when the partition $\Gamma = \Gamma_1 \cup \Gamma_2$ changes, i.e., when the locations of the discontinuities move. However, for a given fixed partition the corresponding system (3.21), (3.34) will be appropriate for any boundary condition of type (3.31), i.e., any combination of the Dirichlet, Neumann, and Robin boundary conditions.

A natural example where the foregoing split of the algorithm into two relatively independent parts may be useful is electromagnetic scattering off conducting materials coated with dielectrics. It is known that the dielectric coating on the surface of a conductor can have a considerable effect on electromagnetic scattering, and even when the shape of the scatterer stays the same and only the coating changes the radar cross section can still vary substantially, see, e.g., [45]. In turn, various types of coating (pure dielectric, lossy dielectric, dielectric with magnetic losses, etc.) can be modeled by the Leontovich [46] or, equivalently, impedance [47, 48] boundary conditions. In the framework of the second order governing equations (Helmholtz-type) those become Robin boundary conditions. The proposed approach will therefore enable an efficient numerical simulation of electromagnetic scattering and radar cross section for a fixed conducting shape that may be fully or partially coated by different types of dielectrics.

3.3 Numerical results

3.3.1 Solution of multiple problems at low cost

The algorithm of the method of difference potentials described in Chapter 3.2.7 allows computationally inexpensive solutions to problems which share certain similar features. The most expensive component of the algorithm is the application of the projection operator, which involves the inverse $\mathbf{G}^{(h)}$ of the finite difference operator $\mathbf{L}^{(h)}$, i.e., the solution of the discrete AP (2.52) by finite differences, see Section 3.1.1. Building the matrix \mathbf{Q}_H of (3.22) that enters into the BEP (3.4) (or, in the inhomogeneous case, (3.5)) requires that the projection operator be applied to each basis function once it has been extended to the discrete boundary γ . In the case $\Gamma = \Gamma_1 \cup \Gamma_2$, in which \mathbf{Q}_H takes the form (3.34), this amounts to a total of $4(N + 1)$ applications of the projection [cf. the number of Chebyshev basis functions in (3.33)].⁵ for the Chebyshev basis. Additionally, for inhomogeneous problems there will be two applications of the inverse operator $\mathbf{G}^{(h)}$ which are associated with the right-hand side f of (3.1a), one for the term $\mathbf{Tr}\mathbf{G}^{(h)}f$, and one more to compute the inhomogeneous contribution to the extension, \mathbf{Q}_I . One final application of the projection is required to obtain the final numerical solution since it

⁵For simplicity, we take $N_1 = N_2 = N$ throughout our simulations. The exception to this is the study of Section 3.3.7, in which we study the effects of taking fewer basis functions to restore accuracy on coarse grids and give more careful consideration to determining a suitable number of basis functions, including the allowance of different numbers of basis functions for each segment of the boundary (i.e., $N_1 \neq N_2$). We also use different numbers of basis functions on each segment for the inhomogeneous problems of Section 3.3.8

is expressed by the generalized Green's formula (3.6) once the system for the basis coefficients has been solved via QR .

Therefore, provided that matrix Q_H is unchanged, the overall computational cost of solving a new problem is small since the projection operator need not be applied again to the basis functions. All of the following simulations are for the Helmholtz equation with constant coefficients, which means that, due to the structure of the chosen AP (2.51), an efficient solution by the separation of variables can be used; however, because our interest is more generally in variable-coefficient equations we implement a direct LU solver for the AP. Alternatively, one may solve the AP with an iterative method. Regardless of the specific solver employed, the algorithm is more efficient when a previously computed basis can be reused since this drastically reduces the number of APs that need to be solved for the new problem. In particular, for problems that differ only in the right-hand sides f in (3.1a), the cost will be that of the 2 projections associated with the right-hand side and 1 more to obtain the final solution. If only the boundary condition (3.1b) is changed, then only the 1 projection to obtain the final solution by the difference potential is needed.

Such savings can be achieved by the design of the numerical examples of Sections 3.3.4-3.3.10, and we demonstrate these savings when applicable.

3.3.2 Parameters of the computational setting

For all of the following test cases, the domain Ω of the interior BVP for the constant coefficient Helmholtz equation (3.1a) is a disk of radius 1 centered at the origin, i.e., the boundary curve Γ is a circle.

Auxiliary Problem The auxiliary domain is a square of side length 2.2 also centered at the origin. The simulations are conducted using the fourth-order accurate compact finite difference scheme (2.33) on a series of Cartesian grids containing 64, 128, 256, 512, 1024, and 2048 cells uniformly spaced in each direction. Scheme (2.33) is supplemented by the Sommerfeld-type boundary conditions (2.50a-2.50b) at the left and right edges of the auxiliary square, and a Dirichlet condition (2.44a) at its top and bottom edges. In particular, note that the circular boundary Γ is not aligned with the Cartesian grid of the FD scheme.

Basis Functions As mentioned in Section 3.2.3, the number of basis functions N used to expand ξ_Γ by formula (3.17) may be chosen grid-independent.⁶ Specifically, it is taken as a num-

⁶While we always use one and the same notation N , it represents $2N + 1$ functions for the Fourier basis, $e^{-iN\theta}, \dots, 1, \dots, e^{iN\theta}$, and N functions for the Chebyshev basis, $T_0(x), \dots, T_{N-1}(x)$. Moreover, as ξ_Γ is a two-component vector function, we need a separate system of basis functions for each component, see (3.17). This makes the overall dimension equal to $2(2N + 1)$ for the Fourier basis and $2N$ for the Chebyshev basis.

ber sufficiently large to represent the given boundary functions α , β , and ϕ , see formulae (3.24) and (3.31), to a prescribed tolerance, 10^{-10} . This boundary tolerance is selected beforehand to exceed any accuracy that we expect to obtain on all grids. When the boundary coefficients and data are continuous and piecewise-smooth, the corresponding Fourier or Chebyshev expansions converge fast on each interval of smoothness, and the resulting number N is relatively small. Altogether, this is a robust and universal strategy that allows us to choose the boundary representation once and for all and then control the final accuracy exclusively by adjusting the grid size. An alternative approach would be to perform a series of computations on the coarsest grids first in order to determine the accuracy needed and then extrapolate out the requirement for finer grids — this approach is given in detail and implemented in Section 4.3. However, for the simulations of this chapter we have chosen to keep the same number of basis functions for all grids as a matter of simplifying the analysis of the computational complexity, as it allows for an unbiased measure of the scaling of the problem relative to the grid size, see Sections 3.3.3–3.3.5.

We realize that this strategy may result in an excessive number of basis functions since the boundary expansion needs to be only as accurate as the finite difference scheme for a particular grid. In other words, for coarser grids one can take fewer basis functions (3.16) and still achieve the greatest accuracy of which the finite difference scheme is capable of on that grid. Moreover, taking too many basis functions on coarse grids may actually result in a loss of accuracy. Indeed, if the dimension of the basis on the boundary is higher than the number of nodes $|\gamma|$, then the least-squares problems derived from (3.21) are no longer overdetermined. Additionally, for the Chebyshev basis functions in particular, the density of the roots of the polynomials near the endpoints can result in a loss of accuracy without the system (3.21) becoming overdetermined. In this case, the oscillations of the basis functions on the continuous boundary Γ are finer than the resolution afforded by the discrete boundary γ . The effect of this phenomenon is that two basis functions may become indistinguishable on the discrete boundary, leading to a degradation of linear independence in the discrete space, and this subsequently introduces error into the least-squares solution of (3.21). We have observed this loss of accuracy on coarse grids in Sections 3.3.3–3.3.5, giving the appearance there of an inflated convergence rate. Therefore, in Section 3.3.6 we demonstrate that reducing the number of basis functions on these coarse grids restores the accuracy of the least-squares solution. This provides for not only more accurate, but also more efficient computation on the coarser grids. Moreover, in Section 3.3.7 we demonstrate that further efficiency can be achieved on every grid in asymmetrically split boundaries $\Gamma = \Gamma_1 \cup \Gamma_2$ by allowing for a different number of basis functions on each segment.

Errors For the examples of Sections 3.3.3–3.3.8, we take $k = 10$ in the Helmholtz equation (3.1a), and consider a known smooth exact solution. In the homogeneous problems of Sections 3.3.3–3.3.7 and 3.3.9, the test solution used will always be a plane wave, while various

test solutions are presented for the inhomogeneous problems of Section 3.3.8, and applying the Helmholtz operator L to the test solutions determines the different right-hand sides f . Subsequently, because the exact solution is known, the error is computed in the maximum norm on the set of nodes \mathbb{N}^+ (see Figure 3.1) and the convergence rate is determined by taking a binary logarithm of the ratio of the errors on successively doubled grids.

We also choose a wavenumber of $k = 5$ in Section 3.3.10 and $k = 10$ in Section 3.3.11. However, for these problems the exact solution is not known. Therefore, in these cases the convergence is assessed by evaluating the maximum norm of the difference between the numerical solutions obtained on pairs of consecutive grids.

Chebyshev endpoints For the examples of Sections 3.3.4–3.3.9, the circle is decomposed into two arcs which meet, by design, at the points of discontinuity of the Robin coefficients, with independent Chebyshev bases on each arc. The same is done at the points of discontinuity in the Dirichlet or Neumann data in Sections 3.3.10–3.3.11. In doing so, the trace of the solution on each arc of the circle is represented separately by a set of Chebyshev basis functions.⁷ To avoid computing the derivatives of the Chebyshev functions near the endpoints, we implement a Chebyshev basis on the extended interval, as described in Section 3.2.6. In doing so, the “gap” ε is estimated as follows. The “worst term” in the extension operator (3.7) applied to a given a Chebyshev basis function on Γ will be that with the highest normal derivative, because it translates into the highest tangential derivative according to (3.12–3.13). The highest tangential derivative is of order four, and it contains the problematic term $(1 - x^2)^{7/2}$ in the denominator⁸ (see Section 3.2.6). Given that the machine precision is on the order of 10^{-16} , we seek ε such that $(1 - x^2)^{7/2} \leq 10^{-10}$ when $|x| > 1 - \varepsilon$ to provide a rough estimate. This yields $\varepsilon > 0.0007$, and thus we have conservatively chosen $\varepsilon = 0.001$ for our computations. This ensures that no calculations of the derivatives of the Chebyshev basis functions will occur within the problematic region near the endpoints $x = \pm 1$. By comparing the results of simulations in Section 3.3.5 (Chebyshev basis on an extended interval with $\varepsilon = 0.001$) with those in Section 3.3.9 (an unmodified Chebyshev system, i.e., with $\varepsilon = 0$), we demonstrate that the approach of Section 3.2.6 indeed provides a very efficient remedy for the “near-singular” behavior of the Chebyshev derivatives at the endpoints.

⁷Those bases could have different dimensions, as we demonstrate in Section 3.3.7. The larger of the two arcs may require more basis functions than the smaller, but we have, for simplicity, used the same number of basis functions N on each arc. In this case, the overall number of basis functions is $4N$, i.e., $2N$ per one arc.

⁸Recall, since Chebyshev functions are polynomials, neither they nor their derivatives can become singular. In other words, the apparent zero in the denominator is canceled by the same order of zero in the numerator. However, this l’Hôpital-type indeterminacy makes numerical computation of the derivatives at the endpoints difficult.

Computational complexity The computer implementation of the entire algorithm is done in MATLAB. To assess the computational complexity, we measure the run time for different parts of the code. The overall computation time is dominated by applying the projection operator (3.3) to the extended basis functions to form the matrix \mathbf{Q} of (3.22), since this requires the expensive step of solving the discrete AP (2.52), i.e., inverting the finite difference operator $\mathbf{L}^{(h)}$. By representing $\mathbf{L}^{(h)}$ as a matrix, this is accomplished using MATLAB's built-in sparse direct solver that employs the LUPQR factorization. The factorization itself is obviously performed only once and then applied to multiple right-hand sides consisting of the extended basis functions. Hence, the cost of building the matrix \mathbf{Q} is split into two parts: the cost of LU factorization and the cost of multiple backward substitutions.

Once the matrix \mathbf{Q} has been formed, the next step is to take into account the boundary condition via the supplemental equations (3.29) or (3.35) and to solve the resulting overdetermined linear system in the sense of least squares. This is done by means of the QR factorization, and we measure the corresponding CPU times for the various cases that we investigate. Finally, once all the coefficients of expansion (3.17) have been determined, one needs to solve the discrete AP (2.52) one more time to obtain the solution $u = \mathbf{P}_{\mathbb{N}^+} \xi_\gamma$, where $\xi_\gamma = \mathbf{E} \mathbf{x} \xi_\Gamma$ is the reconstructed density (for inhomogeneous problems, an additional term is needed along with the potential so that $u = \mathbf{P}_{\mathbb{N}^+} \xi_\gamma + \mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$, see (3.6)). This amounts to performing one additional backward substitution (and one more in the inhomogeneous case for the term $\mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$), as the LUPQR factorization of the matrix $\mathbf{L}^{(h)}$ stays the same.

Altogether, the cost of QR, as well as that of the final solution, is much smaller than the cost of building the matrix \mathbf{Q} . This implies that if \mathbf{Q} is available, changing the boundary condition or right-hand side f of the BVP (3.1) and solving the resulting problem can be done very economically, see Section 3.3.1. Note that the same matrix \mathbf{Q} can be used in Sections 3.3.4-3.3.9, with the technical exception that for the inhomogeneous problems of Section 3.3.8 this \mathbf{Q} corresponds to the homogeneous component \mathbf{Q}_H , and the inhomogeneous component $\mathbf{Q}_I f$ must be formed for each new right-hand side f , which requires one additional backsolve of the matrix $\mathbf{L}^{(h)}$ to obtain. For clarification, note that the \mathbf{Q} -system in the inhomogeneous case (3.23) requires both the terms $\mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$ and $\mathbf{Q}_I f$, which each require one back solve of $\mathbf{L}^{(h)}$. This represents the total of additional computation required for an inhomogeneous problem as opposed to a homogeneous problem since the quantity $\mathbf{G}^{(h)} \mathbf{B}^{(h)} f^{(h)}$ can then be saved and reused for computing the final solution by the potential (3.6) rather than being re-computed. The final set of experiments regarding the effect of the order of the discontinuity on convergence in Sections 3.3.10-3.3.11 has discontinuities at $\theta = 0, \pi$; accordingly, these tests required that the boundary be partitioned differently. This implies that the Chebyshev bases will change, and so a different matrix \mathbf{Q} was computed for the first problem of those simulations and was subsequently reused in all remaining tests.

3.3.3 Smooth periodic Robin boundary condition via Fourier basis

Consider the Robin boundary condition (3.24) with the smooth, periodic coefficients $\alpha(\theta) = e^{\cos \theta}$ and $\beta(\theta) = e^{\sin \theta}$. We “reverse engineer” the data $\phi(\theta)$ for the Robin condition (3.24) by evaluating the known exact solution $u = e^{ikx}$, $k = 10$, and its normal derivative on the boundary. The results presented in Table 3.1 fully corroborate the theoretical design rate of grid convergence for the proposed methodology (fourth order). We also see that the dominant part of the computational cost is indeed the formation of the matrix \mathbf{Q} of (3.22). In Sections 3.3.4 and 3.3.5, we show how one can efficiently solve a range of boundary value problems using the same matrix \mathbf{Q} .

Table 3.1: Grid convergence and execution times for the smooth periodic Robin boundary condition (3.24). $N = 32$ in formula (3.17), and $\psi_n^{(0)}(\theta) = \psi_n^{(1)}(\theta) = e^{in\theta}$ in formula (3.16).

Grid	Error	Conv. rate	Build \mathbf{Q} time	QR time	$\mathbf{P}_{\mathbb{N}+\xi_\gamma}$ time
64×64	$9.44 \cdot 10^{-3}$	-	0.66	0.095	0.0024
128×128	$5.92 \cdot 10^{-4}$	4.00	2.47	0.13	0.012
256×256	$3.69 \cdot 10^{-5}$	4.00	11.13	0.21	0.081
512×512	$2.29 \cdot 10^{-6}$	4.01	55.07	0.38	0.30
1024×1024	$1.44 \cdot 10^{-7}$	3.99	228.35	0.76	1.42
2048×2048	$8.25 \cdot 10^{-9}$	4.12	1193.07	2.52	6.68

3.3.4 Mixed Dirichlet/Neumann boundary conditions using Chebyshev basis

In this section, as well as in Section 3.3.5, we consider a partition of the boundary into two arcs: $\Gamma = R_1 \cup R_2$, where $R_1 = \{0 \leq \theta < 2\pi/3\}$ and $R_2 = \{2\pi/3 \leq \theta < 2\pi\}$. Following Section 3.2.5, we assign a separate Chebyshev basis to R_1 and R_2 . The dimension of the basis is chosen the same for both arcs: $N = 67$. The combination of the partition $\Gamma = R_1 \cup R_2$ and the basis defines the matrix \mathbf{Q} for each grid (via the extension $\mathbf{E}\mathbf{x}$ of each basis function and the application of the projection (3.3)). In Table 3.2, we present the CPU times needed to build this matrix \mathbf{Q} on every grid that we use. In doing so, we distinguish between the time for LUPQR factorization (done once per grid) and the time for $4N$ backward substitutions (see footnote⁷).

We emphasize that the same matrix \mathbf{Q} will be used to solve all boundary value problems in this section, as well as in Section 3.3.5. We also note that instead of Gaussian elimination, we could have used a more efficient direct solver for the AP, based on the separation of variables

Table 3.2: Time to build the matrix \mathbf{Q} of (3.22).

Grid	LUPQR time	$4N$ back solves time
64×64	0.06	1.16
128×128	0.27	4.28
256×256	1.59	20.38
512×512	8.17	91.03
1024×1024	46.5 1	471.11
2048×2048	1516.08	2563.69

and FFT. Nonetheless, we chose to use LU decomposition because it permits an easy extension to variable coefficients and this data serves as a cost analysis for that case.

Let us now specify a mix of the Dirichlet and Neumann conditions on separate parts of the boundary Γ by choosing the Robin coefficients in formula (3.31) to be either 1 or 0. Specifically, we let $\alpha^{(1)}(\theta) = 1$ and $\beta^{(1)}(\theta) = 0$ for $\theta \in R_1$ and $\alpha^{(2)}(\theta) = 0$ and $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. As in Section 3.2.5, we reduce the general linear system (3.35) to its simplified form (3.36-3.37). The exact solution is taken as $u = e^{ikx}$ with $k = 10$, and is used along with its normal derivate to supply the boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$. Table 3.3 shows the grid convergence results and CPU times for QR and for the final solution. As the overall solution remains smooth, the method yields fourth-order accuracy even though the boundary conditions are mixed.

Table 3.3: Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{ikx}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{\mathbf{N}+\xi_\gamma}$ time
64×64	$5.47 \cdot 10^2$	-	0.011	0.0030
128×128	$3.52 \cdot 10^{-1}$	10.60	0.017	0.013
256×256	$1.61 \cdot 10^{-5}$	14.41	0.035	0.065
512×512	$8.10 \cdot 10^{-7}$	4.32	0.070	0.32
1024×1024	$5.26 \cdot 10^{-8}$	3.95	0.17	1.44
2048×2048	$3.05 \cdot 10^{-9}$	4.11	0.46	7.43

Next, we change the boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ by choosing a different exact solution: $u = e^{iky}$ with $k = 10$. The matrix \mathbf{Q} remains unaffected, and the corresponding convergence

and timing results are presented in Table 3.4. By comparing the CPU times in Tables 3.3 and 3.4 with those in Table 3.2 we see that once the matrix \mathbf{Q} is available, changing the boundary condition and solving a new boundary-value problem becomes an inexpensive task.

Table 3.4: Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{iky}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{\mathbb{N}+\xi_\gamma}$ time
64×64	$2.26 \cdot 10^3$	-	0.011	0.0094
128×128	$2.27 \cdot 10^{-1}$	13.05	0.072	0.012
256×256	$9.71 \cdot 10^{-6}$	14.74	0.17	0.071
512×512	$3.56 \cdot 10^{-7}$	4.77	0.12	0.31
1024×1024	$2.26 \cdot 10^{-8}$	3.98	0.16	1.43
2048×2048	$1.29 \cdot 10^{-9}$	4.15	0.43	7.50

3.3.5 Piecewise smooth Robin boundary condition via Chebyshev basis

In this section, we continue to use the same pair of Chebyshev bases and the same matrix \mathbf{Q} as timed in Table 3.2, but apply them to the Robin boundary condition (3.31) with piecewise smooth coefficients. First, we take $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan \theta + 1$ for $\theta \in R_1$, and $\alpha^{(2)}(\theta) = e^{2 \sin \theta}$, $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. It is easily checked that these coefficients exhibit both jump discontinuities and discontinuities in the first derivative at $\theta = 0$ and $\theta = 2\pi/3$. The boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ are again generated by the smooth exact solution $u = e^{ikx}$ with $k = 10$, and the linear system (3.21) for the Chebyshev coefficients of u and $\frac{\partial u}{\partial \mathbf{n}}$ is supplemented by equations (3.35). The data in Table 3.5 corroborate the fourth-order rate of grid convergence.

Again, as in Section 3.3.4, changing the boundary condition does not require a re-computation of \mathbf{Q} and amounts only to doing another least squares solve and another final solve, see Table 3.6. Note that the QR times for the piecewise smooth Robin cases (Tables 3.5 and 3.6) are longer than those for the mixed Dirichlet/Neumann cases (Tables 3.3 and 3.4). This is because in the mixed Dirichlet/Neumann case we are solving a reduced system (3.21), (3.36), (3.37), whereas in the piecewise Robin case we are solving a full system (3.21) supplemented by equations (3.35), which implies a larger dimension.

Altogether, we see that the method provides the design fourth order accuracy for all the test

Table 3.5: Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) with $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan \theta + 1$ and $\alpha^{(2)}(\theta) = e^{2 \sin \theta}$, $\beta^{(2)}(\theta) = 1$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16). The dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$P_{N+\xi_\gamma}$ time
64×64	$1.92 \cdot 10^3$	-	0.49	0.0023
128×128	$1.62 \cdot 10^{-1}$	13.53	0.56	0.012
256×256	$2.36 \cdot 10^{-5}$	12.74	0.80	0.076
512×512	$1.23 \cdot 10^{-6}$	4.26	1.47	0.31
1024×1024	$8.28 \cdot 10^{-8}$	3.89	2.72	1.44
2048×2048	$4.73 \cdot 10^{-9}$	4.13	5.39	7.52

Table 3.6: Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) with $\alpha^{(1)}(\theta) = e^{\sin \theta}$, $\beta^{(1)}(\theta) = (\theta + 3)^2$ and $\alpha^{(2)}(\theta) = \log(\theta + 3)$, $\beta^{(2)}(\theta) = \sqrt{\theta + 3}$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16). The dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$P_{N+\xi_\gamma}$ time
64×64	$1.95 \cdot 10^3$	-	0.58	0.0024
128×128	$4.65 \cdot 10^{-1}$	12.03	0.68	0.012
256×256	$1.81 \cdot 10^{-5}$	14.64	0.89	0.076
512×512	$1.05 \cdot 10^{-6}$	4.11	1.58	0.31
1024×1024	$6.58 \cdot 10^{-8}$	4.00	2.85	1.44
2048×2048	$3.07 \cdot 10^{-9}$	4.15	5.57	7.29

cases that we have investigated (Tables 3.3–3.6), and also that, once the matrix \mathbf{Q} has been pre-computed (Table 3.2), taking a different boundary condition from a rather broad class can be done at a low computational cost.

3.3.6 Reducing the number of basis functions for coarser grids

In this example, we repeat the first case of Section 3.3.5 with the same parameters except that we alter the number of basis functions for each grid. This alleviates the loss of accuracy on the coarser grids (see the errors for the 64×64 and 128×128 grids in Tables 3.3–3.6) and also reduces the overall execution times. Recall that the loss of accuracy on coarser grids occurs because the vertical dimension $|\gamma|$ of the matrix \mathbf{Q} of (3.22) may be smaller than its horizontal

dimension if too many basis functions are taken. In fact, it may occur even in the case where the system is still formally overdetermined, i.e., its vertical dimension exceeds the horizontal dimension, but only slightly. A possible explanation is that when the grid γ is coarse, the feet of the normals dropped from γ to Γ that are used for building the extension \mathbf{Ex} by formulae (3.7–3.8) do not provide a sufficient resolution for the oscillations of the Chebyshev basis functions, which become increasingly oscillatory with N . This effectively implies that two different continuous basis functions may become practically indistinguishable when extended from Γ to γ , which renders the matrix \mathbf{Q} almost rank-deficient. Every time the grid is refined the vertical dimension $|\gamma|$ roughly doubles, the resolution on Γ increases accordingly, and the foregoing adverse phenomenon quickly vanishes. But in the beginning of our sequence of grids, it manifests itself by a higher-than-expected convergence rate as the matrix \mathbf{Q} becomes “taller” but maintains the same width.

We alleviate this issue by choosing N as follows. From our previous computations (Table 3.5) we know that the error for the final grid will be $\sim 3 \times 10^{-9}$. Hence, we replace our initial tolerance that N should approximate α, β , and ϕ within 10^{-10} by 3×10^{-9} , and this determines the value of N for the finest grid. For this problem this yields $N = 62$ as the number of basis functions needed for the finest grid. To maintain a fourth-order convergence rate, the ratio of the errors on successive grids should be 16. Therefore, for each coarsening of the grid we multiply the tolerance for that grid by 16. Using this rule-of-thumb, we estimate a smaller but sufficient number of basis functions. In doing so, the matrix \mathbf{Q} obtained in Section 3.3.4 does not have to be re-computed; we simply drop the columns that correspond to the basis functions that are not included into a reduced set. The results are presented in Table 3.7.

Table 3.7: Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.31) of Table 3.5 with all the same parameters but fewer basis functions on coarser grids.

Grid	N	Error	Conv. rate	Tolerance	QR time
64×64	43	1.28	-	$5.00 \cdot 10^{-3}$	0.13
128×128	48	$9.67 \cdot 10^{-4}$	10.37	$3.10 \cdot 10^{-4}$	0.21
256×256	53	$2.01 \cdot 10^{-5}$	5.59	$1.94 \cdot 10^{-5}$	0.40
512×512	57	$1.23 \cdot 10^{-6}$	4.03	$1.21 \cdot 10^{-6}$	0.84
1024×1024	61	$8.27 \cdot 10^{-8}$	3.89	$7.57 \cdot 10^{-8}$	2.06
2048×2048	65	$4.71 \cdot 10^{-9}$	4.13	$4.73 \cdot 10^{-9}$	4.88

The errors for the coarse grids with 64×64 and 128×128 nodes in Table 3.7 are several orders

of magnitude smaller than their counterparts in Table 3.5, with the errors for the remaining finer grids being almost identical. With no other distinction between this section and the first example of Section 3.3.5, we have demonstrated that taking unnecessarily many basis functions on Γ relative to the grid dimension can cause a loss of accuracy, and that the grid refinement removes this loss of accuracy. Note also that the QR times in Table 3.7 are somewhat shorter than those in Table 3.5. This is because the horizontal dimensions of the corresponding matrices Q are smaller.

We emphasize that the approach for choosing N that we have used in this section is somewhat artificial because it uses an *a posteriori* estimate of the error on the finest grid. What may be done instead is to obtain an *a posteriori* estimate on the coarsest grid first. Once this is done, we can, rather than working backwards by multiplying the error by 16 each iteration from the finest grid down to the coarsest, divide the error of the coarsest grid by a factor of 16 to determine the expected error for finer grids and the appropriate tolerance for the truncation of the basis. Since the computational time required on the coarsest grid will be much less than on the finest, this is an efficient means of determining the number of basis functions and may serve as an alternative to prescribing an initial tolerance. This is the strategy which we have developed and implemented in Section 4.3 and [26] [see the discussion in Section 4.1].

3.3.7 Increased efficiency by using different numbers of basis functions on individual segments of a split boundary

One should note that the number of basis functions required to achieve a given resolution on each piece the grid boundary may differ when any asymmetry exists either in the geometry or in the splitting itself. The problems of this chapter are completely symmetric with respect to the boundary shape, since we are using a disk centered at the origin; however, we have in Sections 3.3.4-3.3.6 (and also in Sections 3.3.8-3.3.10) chosen a splitting that is uneven: Γ_1 is only half the length of Γ_2 , and thus we should expect to need roughly half as many basis functions on Γ_1 to offer the same level of accuracy as on Γ_2 .

Up to this point, we have chosen for simplicity to take the same number of basis functions on each piece of the boundary shape, but the expressions of Section 3.2.5 make no assumption of having the same number of basis functions on each segment of the boundary⁹.

We now repeat the same test of Section 3.3.6, but with the modification that the tolerance is used to determine the appropriate number of basis functions on each segment independently.

⁹In fact, we have even more freedom in the number of basis functions than is described there since the two-component vector may even have different dimensions for each component of the Cauchy data. There may be cases in which this is desirable, such as when the boundary condition is so simple as to require only a known finite series expansion (e.g., zero Dirichlet boundary conditions would not need more than 1 component in the first entry of the basis pair). This is an avenue left to future study.

The same tolerances for the truncation of the Chebyshev series of the known boundary data utilized in Table 3.7 are now used to determine the number of basis functions on each segment of the discrete boundary. As expected, the results, summarized in Table 3.8, verify that the same level of accuracy is obtained with substantially fewer basis functions. Moreover, the errors on the coarsest grids are reduced even further.

Note that we already have available the matrix \mathbf{Q} due to our prior investigations of this same problem in the preceding sections, and so we are able to perform the following computations efficiently by simply removing columns that coincide with the unused basis functions rather than recomputing \mathbf{Q} . Nevertheless, these results demonstrate that, when solving a new problem with concern for efficiency of the overall algorithm, allowing the number of basis functions to be different at each segment of the boundary can result in substantial computational savings. What Table 3.8 fails to quantify is the computational time that could have been saved by not computing the “extra” columns of the matrix \mathbf{Q} when it was initially computed in the example of Section 3.3.4. As noted in Section 3.3.1, the inclusion of each basis function requires a solution of an AP, which amounts to solving a finite difference problem and is the most computationally expensive step of the algorithm. These potential computations savings would be manifest in reducing the number of back solves to build the matrix \mathbf{Q} in Table 3.2 from $4N$ to the smaller sum $2(N_1 + N_2)$. According to the results of Table 3.8, for this particular problem that would mean $2(38 + 65) = 206$ back solves instead of the $4 * 67 = 268$ back solves that were initially done, a savings of about 23%.

Table 3.8: Convergence data for the piecewise smooth Robin boundary condition (3.31) of Tables 3.5 and 3.7 with a split basis. The total number of basis functions used is $2(N_1 + N_2)$

Grid	(N_1, N_2)	Error	Conv. rate	Tolerance	QR time
64×64	(25,43)	$1.17 \cdot 10^{-2}$	-	$5.00 \cdot 10^{-3}$	0.041
128×128	(28,48)	$3.21 \cdot 10^{-4}$	5.19	$3.10 \cdot 10^{-4}$	0.093
256×256	(31,53)	$1.87 \cdot 10^{-5}$	4.10	$1.94 \cdot 10^{-5}$	0.22
512×512	(33,57)	$1.23 \cdot 10^{-6}$	3.93	$1.21 \cdot 10^{-6}$	0.51
1024×1024	(36,61)	$8.24 \cdot 10^{-8}$	3.90	$7.57 \cdot 10^{-8}$	1.31
2048×2048	(38,65)	$4.72 \cdot 10^{-9}$	4.13	$4.73 \cdot 10^{-9}$	3.05

Moreover, a comparison of the number of basis functions required on each arc from the second column of Table 3.8 shows that the number of basis functions needed on the smaller arc Γ_1 , which is half the size of Γ_2 , is about 58% of what is required on Γ_2 - a little more than half.

We may also compare this problem to that of Section 3.3.3, where we found that using a Fourier basis with $N = 32$ - meaning $2(2N + 1) = 130$ basis functions total - was sufficient for the finest grid of Table 3.1. This implies that, while we have now two independent bases on Γ , we do not outright double the number of basis functions required, as we posited in the concluding remarks of Section 3.2.5, but instead the overall number of basis functions required increases by a factor of about 1.6. We leave a more exhaustive comparison on the effect of further subdivision of the boundary shape on the number of basis functions as an avenue for future study.

3.3.8 Inhomogeneous Problems

In Tables 3.9-3.10 below, we choose a test solution u and obtain f by analytically applying the differential operator L of (3.1a). The right-hand side f as well as its derivatives (which, since we assume f is given analytically, we may freely compute) are then used to compute the additional terms of the discrete BEP (3.5), the inhomogeneous contribution to the extension (3.14) which manifests in the linear system for the basis coefficients (3.23), and the inhomogeneous term of the potential (3.6) to compute the final solution. We specify a simple Dirichlet boundary condition along the entire circular boundary.

We have already shown by the analysis in Tables 3.2-3.8 that the solution of problems which differ only in the boundary condition can be computed efficiently as long as the same basis may be used. In the present section, we further demonstrate the flexibility of the method and its efficiency for solving problems with different source terms, i.e., different right-hand sides f of the BVP (3.1a). Even though this boundary condition has no discontinuities we may continue to use the same splitting of the Chebyshev basis into two segments. This highlights the fact that, supposing one wishes to solve problems on the same domain with multiple right-hand sides and with a host of different boundary conditions that may have singularities at certain known locations, there is no need to compute and maintain multiple different configurations of splits for the basis functions. Instead, one and the same splitting which has segments meeting at each of the known potential points of discontinuity can be used for all such problems, regardless of whether discontinuities are present at all, some, or none of those points. As in Section 3.3.7, we choose the number of basis functions on each boundary segment independently with an adequate truncation tolerance for the Chebyshev expansion of both the boundary conditions and right-hand side.

The final column of Tables 3.9-3.10 lists the time required to perform the two projections which pertain only to inhomogeneous problems. We continue to use the same set of Chebyshev basis functions and, having already computed more than enough of them, we simply reduce the number of columns of \mathbf{Q} . Therefore, the total computation time consists of the solution to the \mathbf{Q} system by least squares, two applications of the projection to the right-hand side f of

(3.1a), and the final projection to obtain the solution to the problem (see the final column of Tables 3.3-3.6).

Table 3.9: Inhomogeneous problem with right-hand side generated by the test solution $u = r^2 e^{ikx}$ with Dirichlet boundary conditions.

Grid	(N_1, N_2)	Error	Conv. rate	QR time	RHS Projection time
64×64	(23,41)	$1.45 \cdot 10^{-3}$	-	0.032	0.0091
128×128	(26,45)	$8.94 \cdot 10^{-5}$	4.02	0.073	0.029
256×256	(28,50)	$5.62 \cdot 10^{-6}$	3.99	0.17	0.12
512×512	(31,54)	$3.53 \cdot 10^{-7}$	3.99	0.50	0.49
1024×1024	(34,59)	$2.22 \cdot 10^{-8}$	3.99	1.18	2.20
2048×2048	(36,63)	$1.21 \cdot 10^{-9}$	4.20	2.64	9.58

Table 3.10: Inhomogeneous problem with right-hand side generated by the test solution $u = r^3 e^{ikx}$ with Dirichlet boundary conditions.

Grid	(N_1, N_2)	Error	Conv. rate	QR time	RHS Projection time
64×64	(23,41)	$8.99 \cdot 10^{-4}$	-	0.034	0.028
128×128	(26,45)	$5.59 \cdot 10^{-5}$	4.01	0.074	0.029
256×256	(28,50)	$3.54 \cdot 10^{-6}$	3.98	0.17	0.12
512×512	(31,54)	$2.21 \cdot 10^{-7}$	4.00	0.51	0.49
1024×1024	(34,59)	$1.37 \cdot 10^{-8}$	4.01	1.28	2.26
2048×2048	(36,63)	$1.33 \cdot 10^{-9}$	3.36	2.71	9.49

3.3.9 Loss of accuracy due to non-treatment of Chebyshev endpoints

In Section 3.2.6, we noted that the position of the grid nodes γ relative to the continuous boundary Γ may require computing the Taylor extension (3.7) of the Chebyshev basis functions very close to (or even precisely at) the Chebyshev endpoints, resulting in a loss of accuracy. We demonstrate this phenomenon, as well as the effectiveness of our method in eliminating it, by re-computing the first example of Section 3.3.5. We do so without the use of any technique to correct the poorly conditioned computation of Chebyshev derivatives near the endpoints,

thereby allowing us to observe the disruption of convergence. We again divide the circle into two arcs: $R_1 = [0, 2\pi/3)$ and $R_2 = [2\pi/3, 2\pi)$, and set the same Robin boundary condition (3.31) with the coefficients $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan(\theta) + 1$ for $\theta \in R_1$ and $\alpha^{(2)}(\theta) = e^{2\sin \theta}$, $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. The boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ are still supplied from the exact solution $u = e^{ikx}$ with $k = 10$. The results are presented in Table 3.11.

Table 3.11: Computation of the same case as that in Table 3.5, but with no special treatment of the Chebyshev endpoints. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.16), and the dimension of the Chebyshev basis on each arc is $N=67$.

Grid	Error	Conv. rate
64×64	$2.53 \cdot 10^3$	-
128×128	$3.49 \cdot 10^{-1}$	12.82
256×256	$2.90 \cdot 10^{-4}$	10.23
512×512	$1.37 \cdot 10^{-4}$	1.08
1024×1024	$9.73 \cdot 10^{-6}$	3.82
2048×2048	$2.56 \cdot 10^{-5}$	-1.39

Comparing the results in Table 3.11 to those in Table 3.5, we observe an immediate decline in the convergence rate for all the grids from 256x256 and finer, which culminates in a complete loss of convergence for the finest grid. Since the example of Section 3.3.5 has precisely the same parameters as the current example except that an extended interval for the Chebyshev basis was used in Section 3.3.5, we conclude that the approach of Section 3.2.6 indeed removes the difficulties that would otherwise arise from computing the derivatives of the Chebyshev functions near the endpoints.

3.3.10 Boundary data that lead to solutions with singularities

Discontinuities in the boundary condition will result in singularities of the solution. Because finite difference schemes are only consistent on classical solutions, this will cause the convergence rate to suffer. We design the following experiments to document this decay in the convergence rate as motivation for the singularity subtraction procedure of Chapter 4 which will restore the accuracy of the method in cases where singularities in the solution arise from discontinuous boundary conditions.

In each of the following experiments, we use the same partition of the circle as before: $R_1 = [0, 2\pi/3)$ and $R_2 = [2\pi/3, 2\pi)$. However, we no longer consider a given smooth exact

solution. We rather specify the boundary data independently on each of the two arcs, R_1 or R_2 , allowing for discontinuities at the points $\theta = 0$ and $\theta = 2\pi/3$. This gives rise to near-boundary singularities in the solution. We also choose $k = 5$ in the Helmholtz equation (3.1a).

On each of the two arcs, R_1 or R_2 , we set either a Dirichlet or Neumann boundary condition. The data we specify on R_2 always correspond to a plane wave with $k = 5$ traveling in the x direction. For this wave, $u(x, y) = e^{ikx}$, the Dirichlet data are $u = e^{ikr \cos \theta}$ and the Neumann data are $\frac{\partial u}{\partial n} = ik \cos \theta e^{ikr \cos \theta}$, where $r = 1$ because Γ is a circle of radius 1. The data on R_1 are intentionally specified with a mismatch, and we consider mismatches of three different types: in the amplitude of the wave, in its traveling direction, and in its phase. In the case of a mismatch in the amplitude, the data specified on R_1 correspond to the wave $u(x, y) = 2e^{ikx}$, which yields $u = 2e^{ikr \cos \theta}$ and $\frac{\partial u}{\partial n} = 2ik \cos \theta e^{ikr \cos \theta}$ on Γ in the Dirichlet and Neumann case, respectively. In the case of a mismatch in the traveling direction, we provide two examples: one with the wave that travels in the y direction and the other with the wave that travels at an angle of 1 radian with respect to the positive x axis. For the wave traveling in the y direction, $u(x, y) = e^{iky}$, on the arc R_1 we have $u = e^{ikr \sin \theta}$ and $\frac{\partial u}{\partial n} = ik \sin \theta e^{ikr \sin \theta}$, $r = 1$, for the Dirichlet and Neumann data, respectively. For the wave traveling at an angle of 1 radian, we have $u = e^{ikr \cos(\theta+1)}$ and $\frac{\partial u}{\partial n} = ik \cos(\theta+1) e^{ikr \cos(\theta+1)}$. Finally, for the experiments in which the waves are out of phase, we have chosen a phase shift of 0.7 radians, yielding $u = e^{i(kr \cos(\theta)+0.7)}$ and $\frac{\partial u}{\partial n} = ik \cos(\theta) e^{i(kr \cos(\theta)+0.7)}$ on R_1 . The numerical results for each type of the mismatch subject to different boundary conditions (Dirichlet, Neumann, and mixed) are summarized in Tables 3.12, 3.13, 3.14, and 3.15. For comparison, we have the continuous case in Table 3.16, i.e., the case with no mismatch, in which the plane wave $u(x, y) = e^{ikx}$ with $k = 5$ supplies the data for both R_1 and R_2 .

As the exact solutions to these problems (except the formulation with no mismatch) are not available, we cannot evaluate the error by comparing the numerical solution to the actual solution on the grid. Instead, we introduce a grid-based metric, which compares the numerical solutions on successive grids that have common nodes. Specifically, we structure our Cartesian grids so that each refinement retains all of the nodes of the previous grid, and then compute the maximum norm of the difference between the two successive solutions on the nodes of the coarser grid. Since this measure involves a pair of grids, Tables 3.12–3.16 display the finer of the pair for each resulting error (i.e., the coarsest grid on which we compute is 64×64). Additionally, we found that the error spikes at the nodes of the discrete boundary γ which are closest to the discontinuities at $\theta = 2\pi/3$ and $\theta = 0$, and that the maximum norm when these points are included does not exhibit significant convergence. Therefore, as an additional modification, we compute the maximum error strictly on the interior of the disk or, more precisely, inside the circle of radius 0.8. We also note that changing the maximum norm to l_2 norm makes no substantial difference in the observed convergence.

From Tables 3.12–3.16 we see that the singularities substantially slow down the convergence. In all the cases with mismatches, the rate of grid convergence is at most first order. This behavior is expected because the scheme essentially loses its consistency near the singularity. The relationship between the regularity of the boundary condition and the loss of convergence is further investigated in Section 3.3.11, and the purpose of Chapter 4 is to restore accuracy when the solution lacks regularity at the boundary.

Table 3.12: Grid convergence for boundary data with amplitude mismatch and either Dirichlet, Neumann, or mixed boundary conditions.

	Dirichlet: R1 and R2		Neumann: R1 and R2		Dirichlet: R1, Neumann: R2	
Grid	Error	Rate	Error	Rate	Error	Rate
128×128	$2.60 \cdot 10^{-2}$	-	$2.64 \cdot 10^{-2}$	-	$1.88 \cdot 10^{-2}$	-
256×256	$1.35 \cdot 10^{-2}$	0.95	$1.61 \cdot 10^{-2}$	0.71	$1.08 \cdot 10^{-2}$	0.80
512×512	$9.91 \cdot 10^{-3}$	0.44	$1.00 \cdot 10^{-2}$	0.69	$5.67 \cdot 10^{-3}$	0.93
1024×1024	$4.80 \cdot 10^{-3}$	1.05	$1.73 \cdot 10^{-3}$	2.53	$2.84 \cdot 10^{-3}$	1.00
2048×2048	$2.87 \cdot 10^{-3}$	0.74	$3.62 \cdot 10^{-3}$	-1.06	$1.68 \cdot 10^{-3}$	0.75

Table 3.13: Grid convergence for boundary data with direction mismatch (plane wave traveling in the y direction) and either Dirichlet, Neumann, or mixed boundary conditions.

	Dirichlet: R1 and R2		Neumann: R1 and R2		Dirichlet: R1, Neumann: R2	
Grid	Error	Rate	Error	Rate	Error	Rate
128×128	$1.42 \cdot 10^{-2}$	-	$6.23 \cdot 10^{-2}$	-	$2.32 \cdot 10^{-2}$	-
256×256	$1.67 \cdot 10^{-2}$	-0.24	$2.85 \cdot 10^{-2}$	1.13	$1.44 \cdot 10^{-2}$	0.69
512×512	$1.04 \cdot 10^{-2}$	0.69	$1.67 \cdot 10^{-2}$	0.77	$6.11 \cdot 10^{-3}$	1.24
1024×1024	$5.33 \cdot 10^{-3}$	0.96	$4.68 \cdot 10^{-3}$	1.84	$6.08 \cdot 10^{-3}$	0.01
2048×2048	$2.86 \cdot 10^{-3}$	0.90	$6.92 \cdot 10^{-3}$	-0.56	$2.70 \cdot 10^{-3}$	1.17

Table 3.14: Grid convergence for boundary data with direction mismatch (plane wave traveling at 1 radian with respect to the positive x direction) and either Dirichlet, Neumann, or mixed boundary conditions.

	Dirichlet: R1 and R2		Neumann: R1 and R2		Dirichlet: R1, Neumann: R2	
Grid	Error	Rate	Error	Rate	Error	Rate
128×128	$4.86 \cdot 10^{-2}$	-	$8.22 \cdot 10^{-2}$	-	$2.94 \cdot 10^{-2}$	-
256×256	$2.49 \cdot 10^{-2}$	0.96	$2.70 \cdot 10^{-2}$	1.61	$1.75 \cdot 10^{-2}$	0.75
512×512	$1.90 \cdot 10^{-2}$	0.39	$1.49 \cdot 10^{-2}$	0.86	$8.25 \cdot 10^{-3}$	1.08
1024×1024	$9.25 \cdot 10^{-3}$	1.04	$7.05 \cdot 10^{-3}$	1.08	$6.61 \cdot 10^{-3}$	0.32
2048×2048	$5.17 \cdot 10^{-3}$	0.84	$7.81 \cdot 10^{-3}$	-0.15	$3.08 \cdot 10^{-3}$	1.10

Table 3.15: Grid convergence for boundary data with phase mismatch (waves out of phase by 0.7 radians) and either Dirichlet, Neumann, or mixed boundary conditions.

	Dirichlet: R1 and R2		Neumann: R1 and R2		Dirichlet: R1, Neumann: R2	
Grid	Error	Rate	Error	Rate	Error	Rate
128×128	$1.78 \cdot 10^{-2}$	-	$1.81 \cdot 10^{-2}$	-	$1.27 \cdot 10^{-2}$	-
256×256	$9.24 \cdot 10^{-3}$	0.95	$1.10 \cdot 10^{-2}$	0.71	$7.41 \cdot 10^{-3}$	0.78
512×512	$6.80 \cdot 10^{-3}$	0.44	$6.86 \cdot 10^{-3}$	0.69	$3.89 \cdot 10^{-3}$	0.93
1024×1024	$3.29 \cdot 10^{-3}$	1.05	$1.19 \cdot 10^{-3}$	2.53	$1.94 \cdot 10^{-3}$	1.00
2048×2048	$1.97 \cdot 10^{-3}$	0.74	$2.48 \cdot 10^{-3}$	-1.06	$1.15 \cdot 10^{-3}$	0.75

Table 3.16: Grid convergence for boundary data with no mismatch and either Dirichlet, Neumann, or mixed boundary conditions.

	Dirichlet: R1 and R2		Neumann: R1 and R2		Dirichlet: R1, Neumann: R2	
Grid	Error	Rate	Error	Rate	Error	Rate
128 × 128	$6.62 \cdot 10^{-5}$	-	$1.35 \cdot 10^{-4}$	-	$6.17 \cdot 10^{-5}$	-
256 × 256	$4.18 \cdot 10^{-6}$	3.99	$1.66 \cdot 10^{-5}$	3.02	$4.02 \cdot 10^{-6}$	3.94
512 × 512	$2.62 \cdot 10^{-7}$	3.99	$1.14 \cdot 10^{-7}$	7.19	$2.53 \cdot 10^{-7}$	3.99
1024 × 1024	$1.65 \cdot 10^{-8}$	3.99	$7.10 \cdot 10^{-9}$	4.00	$1.57 \cdot 10^{-8}$	4.01
2048 × 2048	$1.30 \cdot 10^{-9}$	3.66	$5.64 \cdot 10^{-10}$	3.65	$1.24 \cdot 10^{-9}$	3.66

3.3.11 Effect of boundary data with varying degrees of continuity on convergence

The following set of experiments was conducted in order to systematically explore the effects of discontinuous boundary data on convergence of the scheme by solving problems with singularities of various orders (e.g., discontinuous boundary data, then continuous data with a 1st derivative discontinuity, etc.). The singularity subtraction method of Chapter 4 employs an asymptotic expansion of the behavior of the solution near the singular point. Just as we expect the convergence rate to improve as the boundary data become increasingly smooth, we expect that we will also require fewer terms in the asymptotic expansion of the solution near the singular point in Chapter 4 to restore the design convergence rate.

In the experiments of Section 3.3.10, we specified boundary data on separate segments of the boundary which were, individually, the actual data of a test solution, but which were not from the same test solution, resulting in discontinuities at the segment endpoints. In this experiment, we design a series of boundary data with decreasing degrees of continuity at the segment endpoints. Let Γ be split into the segments $\Gamma_1 = [0, \pi)$ and $\Gamma_2 = [\pi, 2\pi)$ with a Chebyshev basis with $N = 67$ on each segment and the wavenumber $k = 10$ in the Helmholtz equation (3.1a). As the exact solutions are not known, the same grid convergence metric employed in Section 3.3.10 is used to compute the errors in Tables 3.17-3.18.

The following boundary data will be specified on Γ_1 and Γ_2 as either Dirichlet data on both parts or Neumann data on both parts, so that a discontinuity of the actual solution or its normal derivative is guaranteed. The first set of boundary data is as follows:

$$\begin{aligned}\phi_1^{(0)} &= \cos \theta, & \theta \in [0, \pi), \\ \phi_2^{(0)} &= -\cos \theta, & \theta \in [\pi, 2\pi),\end{aligned}\tag{3.43}$$

where the superscript denotes the order of discontinuity - i.e., for the data (3.43) the order of discontinuity is said to be of order zero at the points $\theta = 0, \pi$ since the data itself experiences a jump discontinuity at these points. Boundary data with subsequent orders of discontinuity are obtained by the following formulae:

$$\begin{aligned}\phi_1^{(n)} &= \frac{1}{n!} \sin^n \theta, & \theta \in [0, \pi), \\ \phi_2^{(n)} &= -\frac{1}{n!} \sin^n \theta, & \theta \in [\pi, 2\pi),\end{aligned}\tag{3.44}$$

where we let $n = 1, 2, 3$. Observe that the n^{th} set of boundary data is everywhere continuous in its first $n - 1$ derivatives but experiences a jump discontinuity in its n^{th} derivative at the points $\theta = 0, \pi$. In the following tables, the error corresponding to the n^{th} set of boundary data,

$n = 0, 1, 2, 3$, is denoted by e_n , and the corresponding convergence rate by CR_n .

The results of Tables 3.17-3.18 demonstrate clearly that for both Dirichlet and Neumann problems, convergence is quickly restored as the regularity of the boundary data increases, with steady fourth order convergence being achieved once the data have at least two continuous derivatives. The numerical examples of Section 4.3 follow a similar progression from being less regular to more regular, and we observe in that chapter that fewer singular terms must be accounted for as the regularity of the problem increases. As regularity at the boundary is restored in Tables 3.17-3.18, so is the consistency of the finite difference scheme and thus the convergence rate.

Table 3.17: Error and convergence rate for Dirichlet boundary conditions with 0^{th} up to 3^{rd} order discontinuity using boundary data (3.43-3.44)

Grid	e_0	CR_0	e_1	CR_1	e_2	CR_2	e_3	CR_3
128	30.96	-	1.05	-	61.09	-	$3.31 \cdot 10^{-2}$	-
256	0.298	6.70	$2.46 \cdot 10^{-3}$	8.73	$2.20 \cdot 10^{-4}$	18.08	$1.04 \cdot 10^{-5}$	11.64
512	0.236	0.33	$7.50 \cdot 10^{-4}$	1.71	$1.39 \cdot 10^{-5}$	3.98	$8.24 \cdot 10^{-7}$	3.65
1024	0.367	-0.64	$7.81 \cdot 10^{-4}$	-0.06	$8.74 \cdot 10^{-7}$	3.99	$5.15 \cdot 10^{-8}$	4.00
2048	0.224	0.71	$1.43 \cdot 10^{-4}$	2.45	$2.48 \cdot 10^{-7}$	1.82	$3.24 \cdot 10^{-9}$	3.99

Table 3.18: Error and convergence rate for Neumann boundary conditions with 0^{th} up to 3^{rd} order discontinuity using boundary data (3.43-3.44)

Grid	e_0	CR_0	e_1	CR_1	e_2	CR_2	e_3	CR_3
128	1.41	-	2.72	-	0.528	-	0.534	-
256	$3.72 \cdot 10^{-2}$	5.24	$1.25 \cdot 10^{-3}$	11.08	$5.82 \cdot 10^{-6}$	16.47	$2.28 \cdot 10^{-4}$	11.20
512	$1.90 \cdot 10^{-2}$	0.97	$9.43 \cdot 10^{-5}$	3.73	$7.85 \cdot 10^{-8}$	6.21	$1.45 \cdot 10^{-5}$	3.97
1024	$9.55 \cdot 10^{-3}$	0.99	$1.62 \cdot 10^{-5}$	2.54	$1.11 \cdot 10^{-8}$	2.82	$8.95 \cdot 10^{-7}$	4.02
2048	$4.76 \cdot 10^{-3}$	1.00	$3.91 \cdot 10^{-6}$	2.05	$2.01 \cdot 10^{-9}$	2.47	$5.65 \cdot 10^{-8}$	3.99

The results of Tables 3.17-3.18 show that convergence is restored when the solution to the Helmholtz equation (3.1) becomes classical, i.e., has 2 continuous derivatives. Given that the

FD scheme (2.33) is 4th order, one may expect that 4 continuous derivatives of the solution to the BVP (3.1) would be required to restore convergence. In fact, the problem which we solve by FD is actually the AP (2.51), not the BVP (3.1). Therefore, it is sufficient that solution to the AP (2.51) have 4 continuous derivatives in order to maintain the 4th order design rate of the FD scheme (2.33). In the method of difference potentials, the discontinuity at the boundary curve Γ of the BVP (3.1) manifests itself as the right-hand side g of the AP (2.51). Since the boundary conditions of the AP (2.51) are unchanged by the boundary conditions of the interior BVP (3.1), the right-hand side g is the only source of discontinuity in the AP (2.51). Therefore, if the boundary data of the BVP (3.1) have 2 continuous derivatives, then the right-hand side g of the AP (2.51) will have 2 continuous derivatives, and the solution of the AP (2.51) will have 4 continuous derivatives.

3.4 Discussion and Remarks

We have investigated theoretically and demonstrated experimentally the capability of the method of difference potentials to handle complex boundary conditions, such as variable coefficient Robin, mixed, and discontinuous. The governing Helmholtz equation was approximated on a regular Cartesian grid by an economical fourth order accurate compact finite difference scheme. For a number of test cases that involved a non-conforming circular boundary and various boundary conditions, we have been able to recover the design fourth-order accuracy of the scheme provided that the overall solution was sufficiently smooth. The accuracy was not adversely affected by either staircasing [4, 5] or the non-standard nature of the boundary conditions.

The approach that we use to reduce the original problem from its domain to the boundary is based on Calderon's operators. It automatically guarantees the well-posedness of the resulting boundary formulation as long as the original problem is well posed, regardless of the type of the boundary condition. Moreover, it is very important that when changing either the boundary condition or the right-hand side f of the Helmholtz equation only a particular component of the overall numerical algorithm changes, whereas most of it remains unaffected (see Section 3.2.7). Accordingly, the computational cost associated with solving a new problem for a new boundary condition or with a different source term is small provided that the same basis may be used. In contrast, a change of the boundary condition in the classical method of boundary integral equations often requires a complete change of the algorithm. We note that there is another group of techniques based on finite differences/volumes as opposed to integral equations, and designed to handle non-aligned boundaries/interfaces: immersed boundary [49], immersed interface [50], ghost fluid [51, 52], and embedded boundary [53, 54] methods. To the best of our knowledge, there are no reported uses in the literature of those methods for anything but simple Dirichlet,

Neumann, or interface conditions (continuity of the solution and its normal flux), changing the boundary condition requires major changes to the algorithm [54], and extension to higher than second order accuracy is not straightforward.

We have also shown that when the overall solution is not smooth and has a singularity at the location where the coefficients and/or data in the boundary conditions are discontinuous, then the convergence of the method slows down, as expected, because the finite difference scheme loses its consistency. As the regularity of the boundary data increases, consistency of the finite difference scheme is gradually restored. In Chapter 4, we circumvent the loss of consistency for problems with singular boundary data by first modifying the problem so that the boundary data are sufficiently regular before solving it by the method of difference potentials. The locations of the singular points along the boundary for these problems are known, and these singular points naturally become the endpoints for the partitioning of the boundary curve to ensure that the basis representations of the modified boundary data converge rapidly.

Chapter 4

Singularity Subtraction

As we have seen, the near-boundary singularities of the solution due to discontinuities in the boundary conditions result in a substantial reduction of accuracy for the method of Chapter 3. It is well known that the solution may also develop a singularity due to the geometric irregularities of the boundary itself, e.g., domains with corners. In many cases, the type of the singularity is known, at least to the leading order, see, e.g., [55, 56, 57, 58, 19, 59, 60, 61, 62, 63, 64], as well as the monograph [65]. In such cases, the singular component can be subtracted, and the problem can be solved with respect to the remaining part of the solution. The latter should have higher regularity, and thus enable an improved convergence of the numerical approximation. Combined with the method of difference potentials, this approach was previously implemented for the Laplace equation in [21] and the Chaplygin equation in [20]. None of these efforts used high-order computational methods.

In this chapter, we show how the design accuracy for the combined methodology of Chapter 3 can be restored when the solution has such near-boundary singularities [26]. The idea is to subtract out several leading terms in the expansion of the solution near the singular point, and subsequently solve only for the remaining regular component. In developing the expansion, we follow the approach earlier proposed in [19] and modify it to accommodate singular points on curved boundaries (rather than only straight lines) and to maintain a high-order of accuracy. Following this procedure provides a consistent approach because the resulting system of ODEs for the coefficients of the singular expansion will always have the same structure regardless of the specific geometry of the problem.

For simplicity, the boundary condition (3.1b) will be either of Dirichlet ($\mathbf{l}_\Gamma = 1$) or Neumann ($\mathbf{l}_\Gamma = \frac{\partial}{\partial \mathbf{n}}$) type. The methods presented hereafter extend trivially to the case where (3.1b) is of a mixed type (Dirichlet/Neumann). The boundary data in (3.1b) are now intentionally chosen to be discontinuous. Namely, the function $\phi_\Gamma = \phi_\Gamma(s)$, where s is the arc length, will have either jump discontinuities in the function itself or else jump discontinuities in its first or second

derivative with respect to s .

The aforementioned discontinuities in ϕ_Γ or its derivatives will cause the solution u to possess singularities at the corresponding boundary points. The analysis of Section 4.1 allows us to obtain an approximation, u_S , to those near-boundary singularities (to several leading orders). Then, by substitution into the original homogeneous BVP (3.1), $u_R = u - u_S$ will satisfy the following inhomogeneous BVP:

$$\mathbf{L}u_R = -\mathbf{L}u_S \stackrel{\text{def}}{=} f, \quad \mathbf{x} \in \Omega, \quad (4.1a)$$

$$\mathbf{l}_\Gamma u_R = \phi_\Gamma - \mathbf{l}_\Gamma u_S \stackrel{\text{def}}{=} \psi_\Gamma, \quad \mathbf{x} \in \Gamma. \quad (4.1b)$$

We refer to the new BVP (4.1) as the “regularized” problem since it will now be regular up to the degree with which u_S matches the boundary singularities of u .

For simplicity, we consider the case where the domain Ω is a disk of radius $R=1$ centered at the origin. In particular, we will seek to restore accuracy to problems similar to the numerical study of Section 3.3.11 which have singularities of varying derivative orders at the points $\theta = 0, \pi$. From the standpoint of treating singularities that result from discontinuous boundary data, this choice presents no loss of generality. The regularized BVP (4.1) will have sufficient regularity that it can be solved without loss of accuracy by the method of difference potentials described in Chapter 3.

4.1 Constructing Singular Functions at the Boundary

In this section, we will show how to transition from the original BVP (3.1), for which the solution u may have singularities, to the new BVP (4.1), for which the solution is sufficiently smooth so that it can be approximated by means of a high-order accurate scheme. For definiteness, we will first analyze the case of a Dirichlet boundary condition (3.1b) in problem (3.1), $\mathbf{l}_\Gamma = 1$. The Neumann and mixed cases are treated similarly, and in Sections 4.2 and 4.3 we present the setup and the results of computations for both Dirichlet and Neumann boundary conditions.

We allow the function $\phi_\Gamma(s)$ on the right-hand side of (3.1b) and/or its derivatives to have jump discontinuities at some locations on the boundary Γ . Though it is not the focus of the present work, it is of note that these methods can similarly treat singularities resulting from corners in the domain, allowing for the treatment of boundary shapes that are only piecewise smooth. Discontinuities of either type generally reduce the regularity of the solution in the vicinity of the corresponding boundary points, as some derivatives of the solution become unbounded. This, in turn, slows down the convergence of finite difference methods, see Tables 3.12-3.18. Hereafter, we describe a consistent and general approach to restoring the rate of

convergence affected by the reduced regularity of the solution. It exploits an analytically-derived asymptotic expansion of the solution near the “irregular” boundary points.

Specifically, let u_0, u_1, \dots, u_n be the first $n+1$ consecutive terms of an asymptotic expansion of the exact solution u to the BVP (3.1) in the vicinity of a singular boundary point $s = s_0$. Suppose that each term u_j of this expansion is more regular than the previous term, u_{j-1} , i.e., that u_j has bounded derivatives up to a certain order which is higher than the order of the highest bounded derivative of u_{j-1} . Subtracting the sum of $n+1$ such terms from the exact solution u , we arrive at a new function $v = u - (u_0 + u_1 + \dots + u_n)$, for which we formulate a new BVP [cf. BVP (4.1)]:

$$\begin{cases} \mathbf{L}v = -\mathbf{L}(u_0 + u_1 + \dots + u_n), \\ v|_{\Gamma} = \phi_{\Gamma}(s) - (u_0 + u_1 + \dots + u_n)|_{\Gamma}. \end{cases} \quad (4.2)$$

In general, the right-hand side $-\mathbf{L}(u_0 + u_1 + \dots + u_n)$ in (4.2) is nonzero because neither the individual terms u_j of the expansion nor their sum are expected to satisfy the homogeneous Helmholtz equation. Our goal is to choose the terms u_j in such a way that the function v will have no singularities up to at least the derivative of order n , in which case problem (4.2) will be referred to as regularized.

By taking an appropriate value of n , one can make the solution v of the regularized problem (4.2) sufficiently smooth so that a given finite difference scheme will converge to v with the design rate. Then, by adding the analytically-derived sum $(u_0 + u_1 + \dots + u_n)$ to the computed solution v , one would restore the design order of accuracy for the overall discrete approximation of u .

In the remainder of this section, we derive the asymptotic form of the solution to the Helmholtz equation near the singular points at the boundary. Specifically, in Section 4.1.1 we follow [19] and solve equation (3.1a) near the vertex of a 2D plane wedge with straight boundaries. We focus on the Dirichlet case and describe the Neumann case more briefly. In Section 4.1.2, we use a conformal mapping to generalize the results of Section 4.1.1 for the case of boundaries with nonzero curvature, which permits high-order approximation. As an example, in Section 4.1.3 and Appendix B we consider both Dirichlet and Neumann data when Γ is a unit circle, although our techniques can accommodate more complex geometries as well.

4.1.1 Straight boundaries

Consider the Helmholtz equation (3.1a) in polar coordinates centered at the vertex of a 2D wedge with straight sides and angle ω . Let ρ and θ denote the polar radius and polar angle,

respectively. Then, equation (3.1a) takes the form

$$\left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \theta^2} + k^2(\rho, \theta) \right) u(\rho, \theta) = 0. \quad (4.3)$$

Note though that unlike in (3.1a), we allow for the spatial variation of the wavenumber k in equation (4.3), because it proves useful for the analysis of curved boundaries in Section 4.1.2. Indeed, after the conformal mapping that we use to straighten the boundaries, the constant-coefficient Helmholtz equation transforms into a variable-coefficient Helmholtz equation. Thus, we must necessarily treat the variable-coefficient case.

Equation (4.3) is supplemented by the boundary conditions on each side of the wedge:

$$u(\rho, 0) = F(\rho), \quad u(\rho, \omega) = H(\rho). \quad (4.4)$$

We assume that the data in (4.4) can be expanded into the power series¹:

$$F(\rho) = \sum_{j=0}^{\infty} f_j \rho^j, \quad H(\rho) = \sum_{j=0}^{\infty} h_j \rho^j. \quad (4.5)$$

In doing so, expansions (4.5) are not required to match at the vertex of the wedge. In other words, the respective coefficients f_j and h_j in (4.5) may differ from one another. It is precisely this mismatch between the boundary data that will give rise to the singularities in the solution.

In addition, we assume that a convergent expansion exists for $k^2(\rho, \theta)$ as well:

$$k^2(\rho, \theta) = \sum_{j=0}^{\infty} k_j(\theta) \rho^j. \quad (4.6)$$

Representations (4.5), (4.6) for the boundary data and the wavenumber, respectively, will be needed for constructing the asymptotic expansion of the solution $u = u(\rho, \theta)$ near the singularity. Following [19], we seek an asymptotic expansion of the solution to (4.3) in the form of a series

$$u(\rho, \theta) = \sum_{j=0}^{\infty} \rho^j (A_j(\theta) \ln \rho + B_j(\theta)). \quad (4.7)$$

Substituting expansions (4.6) and (4.7) into equation (4.3) and requiring that the resulting coefficients in front of all terms ρ^j and $\rho^j \ln \rho$, $j = 0, 1, 2, \dots$, be independently equal to zero,

¹We have not deemed it necessary for this study, but it may in some cases be convenient to allow the power series to have a more general form in which the exponent of r is fractional, as in the work of Fox and Sankar [19]. This does not affect the ODE systems for A_j and B_j given in equations (4.8) and (4.9)

we obtain two systems of second order ODEs for $A_j(\theta)$ and $B_j(\theta)$:

$$\begin{cases} A_0'' = 0, \\ A_1'' + 1^2 \cdot A_1 = 0, \\ \vdots \\ A_{m+2}'' + (m+2)^2 A_{m+2} = - \sum_{j=0}^m k_{m-j} A_j, \quad m = 0, 1, 2, \dots, \\ A_j(0) = A_j(\omega) = 0, \quad j = 0, 1, 2, \dots, \end{cases} \quad (4.8)$$

$$\begin{cases} B_0'' = 0, \\ B_1'' + 1^2 \cdot B_1 = -2A_1, \\ \vdots \\ B_{m+2}'' + (m+2)^2 B_{m+2} = - \sum_{j=0}^m k_{m-j} B_j - 2(m+2)A_{m+2}, \quad m = 0, 1, 2, \dots, \\ B_j(0) = f_j, \quad B_j(\omega) = h_j, \quad j = 0, 1, 2, \dots \end{cases} \quad (4.9)$$

In formulae (4.8) and (4.9), primes denote differentiation with respect to θ , and the boundary conditions for $B_j(\theta)$ at $\theta = 0$ and $\theta = \omega$ in (4.9) are obtained with the help of expansions (4.5).

Systems (4.8) and (4.9) are coupled and should therefore be solved concurrently, with individual equations addressed in the consecutive order starting with those for $A_0(\theta)$, $B_0(\theta)$. Schematically, we represent the solution sequence as follows: $A_0(\theta) \rightarrow B_0(\theta) \rightarrow A_1(\theta) \rightarrow B_1(\theta) \rightarrow A_2(\theta) \rightarrow B_2(\theta) \rightarrow \dots$. At each step of this sequence, $j = 0, 1, 2, \dots$, we solve an ODE of the type

$$\begin{cases} \Psi_j'' + j^2 \Psi_j = \mu_j(\theta), \\ \Psi_j(0) = a_j, \quad \Psi_j(\omega) = b_j. \end{cases} \quad (4.10)$$

In the case that $\Psi_j = A_j$, the boundary conditions are the homogeneous conditions of (4.8) (i.e., $a_j = b_j = 0$), and (4.10) becomes a Sturm-Liouville problem. In this case, the right-hand side $\mu_j(\theta)$ in (4.10) involves lower-order coefficients $A_i(\theta)$, $i < j$, in the form of a convolution with the coefficients $k_j(\theta)$. Likewise, when $\Psi_j = B_j$ the boundary conditions are given by $a_j = f_j$ and $b_j = h_j$ from (4.9), and the right-hand side $\mu_j(\theta)$ includes both lower order coefficients $B_i(\theta)$, $i < j$, and the coefficient $A_j(\theta)$ of the same order.

We will now study the solvability of problem (4.10). To do so, it will be sufficient to consider only the Sturm-Liouville problem with $a_j = b_j = 0$. Indeed, for system (4.8) the boundary conditions are homogeneous anyway, and for system (4.9) they can be easily made homogeneous

by subtracting an arbitrary (smooth) function $\tilde{\Psi}_j(\theta)$ from the solution such that $\tilde{\Psi}_j(0) = a_j$, $\tilde{\Psi}_j(\omega) = b_j$, and adjusting the right-hand side accordingly. Hence, the reduction of problem (4.10) to the case $a_j = b_j = 0$ presents no loss of generality. At the same time, it considerably simplifies the analysis and discussion below.

The Fredholm alternative holds for the Sturm-Liouville problem (4.10) for any given j and $a_j = b_j = 0$. It means that one of the following two scenarios transpires:

1. Problem (4.10) has a unique solution, which takes place if $\omega \neq \pi \frac{l}{j}$, $l = 1, 2, \dots$. This is a non-resonant case: the Sturm-Liouville operator $L_j \equiv \frac{d^2}{d\theta^2} + j^2$ subject to zero Dirichlet conditions will not resonate since the second derivative operator has no non-trivial eigenfunctions; the inverse operator (Green's function) L_j^{-1} exists; and the problem $L_j \Psi_j = \mu_j$, $\Psi_j(0) = \Psi_j(\omega) = 0$, can be uniquely solved.
2. The resonant case. At $\omega = \pi \frac{l}{j}$, $l = 1, 2, \dots$, there exists a nonzero eigenfunction of the second derivative operator of the Sturm-Liouville operator L_j . This eigenfunction is merely $\hat{\Psi}_j(\theta) = \sin(j\theta)$, up to a constant factor. In this case, problem (4.10) has a solution if and only if the solvability condition holds²:

$$\int_0^\omega \mu_j(\theta) \sin j\theta d\theta = 0. \quad (4.11)$$

The solution is given by

$$\Psi_j(\theta) = C_j \sin j\theta + \frac{1}{j} \int_0^\theta \{\cos j\theta' \sin j\theta - \sin j\theta' \cos j\theta\} \mu_j(\theta') d\theta', \quad (4.12)$$

and is not unique, since the coefficient C_j in front of the eigenfunction $\sin j\theta$ is arbitrary and changing it will violate neither the differential equation nor the zero boundary conditions.

Next, it will be convenient to represent the wedge angle as $\omega = \alpha\pi$, where α may be either a rational number or an irrational number. Let us first assume that α is irrational. Then, problem (4.10) falls into the first proposition of the Fredholm alternative. The coefficients $A_j(\theta)$ are all identically equal to zero, $A_j(\theta) \equiv 0$, $j = 0, 1, 2, \dots$, since the boundary conditions are zero and the solution is unique. The coefficients $B_j(\theta)$ are uniquely determined and in general are

²To verify this, consider the following chain of equalities: $0 = \int_0^\omega \Psi(\theta) L_j \hat{\Psi}(\theta) d\theta = \int_0^\omega \hat{\Psi}(\theta) L_j \Psi(\theta) d\theta = \int_0^\omega \hat{\Psi}(\theta) \mu_j(\theta) d\theta$, where we have used the integration by parts and taken into account that $L_j \hat{\Psi}(\theta) = 0$ and $L_j \Psi_j = \mu_j$.

not equal to zero identically due to the nonzero boundary conditions. Consequently, for the angles $\omega = \alpha\pi$ with α irrational, we obtain a regular expansion of the solution with no singular logarithmic terms.

If α is a rational number, i.e., $\alpha = \frac{l}{N}$, where l and N are positive integers with a greatest common factor of 1, then for those j_k for which $j_k \frac{l}{N}$ is also a positive integer,

$$j_k \frac{l}{N} = c_k \in \mathcal{N}, \quad (4.13)$$

problem (4.10) falls into the second proposition of the Fredholm alternative. For these j_k , the coefficients $A_{j_k}(\theta) = C_{j_k}^{(A)} \sin j_k \theta$ are eigenfunctions of L_{j_k} , up to an arbitrary factor $C_{j_k}^{(A)}$. Substituting a given $A_{j_k}(\theta)$ into the right-hand side of the equation (4.9) for $B_{j_k}(\theta)$ and applying the solvability condition (4.11), we obtain an algebraic equation for the constant $C_{j_k}^{(A)}$ that can be uniquely solved.

On the other hand, the coefficient $B_{j_k}(\theta)$ appears only partially determined. While the second term in (4.12) is defined unambiguously, the first term contains the eigenfunction $\sin j_k \theta$ with an arbitrary constant C_{j_k} in front of it. Therefore, in the case of a rational α we obtain nonzero, uniquely determined coefficients $A_{j_k}(\theta)$ for the singular terms $\rho^{j_k} \ln \rho$, and partially determined coefficients $B_{j_k}(\theta)$ for ρ^{j_k} . The successive solution of the coupled systems (4.8), (4.9) can be schematically shown as

$$A_0(\theta) \rightarrow B_0(\theta) \rightarrow \dots \rightarrow A_j(\theta) \rightarrow B_j(\theta) \rightarrow \dots \rightarrow A_{j_k}(\theta) \rightarrow B_{j_k}(\theta) \rightarrow \dots \quad (4.14)$$

$\begin{array}{c} \uparrow \qquad \qquad \downarrow \\ \longleftarrow \\ C_{j_k}^{(A)} \end{array}$

Diagram (4.14) emphasizes that at each step of the sequence the problem falls into one of the two propositions of the Fredholm alternative and either proceeds straightforwardly in the non-resonant case (like the j -th step in (4.14)) or makes a back loop in the resonant case (like the j_k -th step).

We can also see from (4.13) that the first non-zero singular term $A_{j_k}(\theta) \rho^{j_k} \ln \rho$ appears in the series at $j_1 = N$. So, as N increases the expansion becomes more regular because the singularity moves to higher-order terms. In the limit $N \rightarrow \infty$, $l \rightarrow \infty$, the ratio l/N approaches an irrational number as long as it remains finite and irreducible. Therefore, the singularity disappears and we arrive at the case of an irrational α discussed above.

In the case of a Neumann problem, the boundary conditions are set for the normal derivatives on the sides of the wedge. For the previously described geometry, this is the same as specifying

$\frac{1}{\rho} \frac{\partial u}{\partial \theta}$ at $\theta = 0$ and $\theta = \omega$:

$$\frac{1}{\rho} \frac{\partial u(\rho, 0)}{\partial \theta} = F(\rho) = \sum_{j=0}^{\infty} f_j \rho^j, \quad \frac{1}{\rho} \frac{\partial u(\rho, \omega)}{\partial \theta} = H(\rho) = \sum_{j=0}^{\infty} h_j \rho^j,$$

or, equivalently,

$$\frac{\partial u(\rho, 0)}{\partial \theta} = \rho F(\rho) = \sum_{j=0}^{\infty} f_j \rho^{j+1}, \quad \frac{\partial u(\rho, \omega)}{\partial \theta} = \rho H(\rho) = \sum_{j=0}^{\infty} h_j \rho^{j+1}. \quad (4.15)$$

Expansions (4.15) suggest that one can seek a solution in the form of a series

$$u(\rho, \theta) = \sum_{j=0}^{\infty} \rho^{j+1} (A_j(\theta) \ln \rho + B_j(\theta)) \quad (4.16)$$

which differs from (4.7) only in that the respective powers of ρ are increased by one. Expansion (4.16) yields the systems similar to (4.8-4.10), but with different coefficients in the equations and with Neumann boundary conditions in the counterpart of the Sturm-Liouville problem (4.10). The procedure for solving these coupled systems (for $A_j(\theta)$ and $B_j(\theta)$) remains the same as for the Dirichlet case.

We can also consider a more general class of the boundary data, with expansions that include non-integer powers:

$$F(\rho) = \rho^\beta \sum_{j=0}^{\infty} f_j \rho^j, \quad H(\rho) = \rho^\gamma \sum_{j=0}^{\infty} h_j \rho^j, \quad 0 < \beta, \gamma < 1.$$

In this case the overall problem should be split into two subproblems using linear superposition. The first subproblem will have the expansion

$$u^{(1)}(\rho, \theta) = \sum_{j=0}^{\infty} \rho^{j+\beta} (A_j(\theta) \ln \rho + B_j(\theta)),$$

with $F^{(1)}(\rho) = F(\rho)$ and $H^{(1)}(\rho) = 0$ as the boundary data. Likewise, the second expansion will have the form

$$u^{(2)}(\rho, \theta) = \sum_{j=0}^{\infty} \rho^{j+\gamma} (A_j(\theta) \ln \rho + B_j(\theta)),$$

with $F^{(2)}(\rho) = 0$ and $H^{(2)}(\rho) = H(\rho)$, respectively.

4.1.2 Curved boundaries

In this section, we build an asymptotic expansion of the solution to the Helmholtz equation (3.1a):

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)u(x, y) + k^2 u(x, y) = 0 \quad (4.17)$$

near a point on the curved (i.e., non-straight) boundary where the data undergo a discontinuity. We assume that the domain Ω is a disk of radius R centered at the origin on the (x, y) -plane.³ Introducing the polar coordinates (r, φ) , we can write the boundary condition (3.1b) as follows:

$$u|_{r=R} = \phi_\Gamma(\varphi). \quad (4.18)$$

The function $\phi_\Gamma(\varphi)$ and/or its derivatives may have jump discontinuities for some given values of φ . Furthermore, the wavenumber k in equation (4.17) may, in general, vary with coordinates, $k = k(x, y)$, and we have chosen a constant wavenumber only for simplicity, since, regardless, it becomes a variable quantity after the conformal mapping that we introduce below.

Let $\varphi = 0$ be one of the points on the boundary where $\phi_\Gamma(\varphi)$ and/or its derivatives is discontinuous.⁴ We build an asymptotic expansion to the solution of equation (4.17) in the vicinity of $(R, 0)$. For that purpose, we use a conformal mapping $z = x + iy \mapsto \zeta = \xi + i\eta$ that reduces the problem to the case we have analyzed previously, in Section 4.1.1. The fractional linear transform

$$\zeta = i \frac{R - z}{R + z} \quad (4.19)$$

maps the disk of radius R onto the half-plane $\eta \geq 0$ in the (ξ, η) -coordinates (see Figure 4.1). The upper (lower) semi-circle is mapped onto the positive (negative) real semi-axis $\xi > 0$ ($\xi < 0$). The point $(R, 0)$ is mapped into the origin $(0, 0)$ of the (ξ, η) -plane whereas the opposite point $(-R, 0)$ corresponds to $\pm\infty$ on the real axis ξ .

It is well known that under a general conformal mapping $\zeta = \zeta(z)$, the Helmholtz equation (4.17) transforms as follows:

$$\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right)u(\xi, \eta) + |z'(\zeta)|^2 k^2 u(\xi, \eta) = 0, \quad (4.20)$$

where $z = z(\zeta)$ is the function inverse to $\zeta = \zeta(z)$, and where a prime denotes differentiation

³Note that we have chosen $R = 1$ for simplicity; this will be our choice for the numerical simulations of Section 4.3 as well. Moreover, the choice of a circular shape in the first place is also only a matter of convenience as it makes the conformal mapping that we are going to use particularly simple (fractional linear). Otherwise, it does not present a loss of generality, and our technique can, in fact, address a singularity at the vertex of an arbitrary wedge with curved sides.

⁴This assumption involves no loss of generality as we can always rotate the frame.

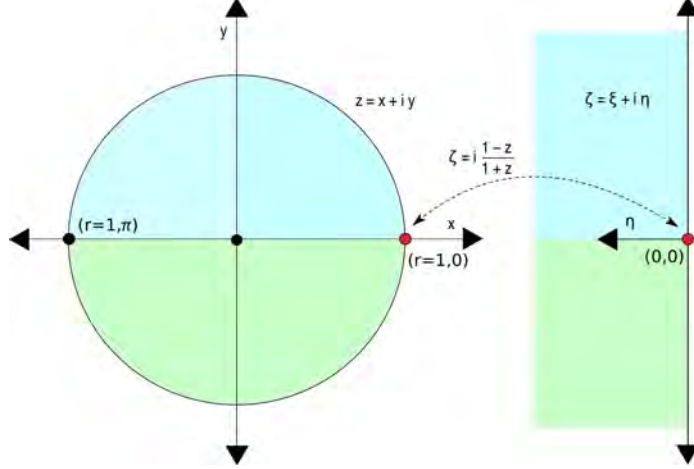


Figure 4.1: Conformal mapping (4.19).

(in the Cauchy-Riemann sense). For mapping (4.19), equation (4.20) becomes:

$$\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} \right) u(\xi, \eta) + \frac{4k^2 R^2}{[(\xi + 1)^2 + \eta^2]^2} u(\xi, \eta) = 0. \quad (4.21)$$

We observe that now the wavenumber varies as a function of ξ and η . The boundary condition (4.18) for equation (4.21) is set on the real axis ξ (i.e., the axis $\eta = 0$):

$$u(\xi, 0) = f(\xi), \quad (4.22)$$

where the function $f(\cdot)$ is obtained from the original $\phi(\cdot)$ of (4.18) via the inverse transform $z = z(\zeta)$. Any discontinuities of the boundary data at the point $\varphi = 0$ in (4.18) are therefore translated to the point $\xi = 0$ after the mapping (4.19).

Next, we introduce polar coordinates $\xi = \rho \cos \theta$, $\eta = \rho \sin \theta$ on the (ξ, η) -plane, see Figure 4.1, and recast equation (4.21) on the half-plane $\eta \geq 0$ in terms of ρ and θ . This yields the Helmholtz equation in the form (4.3) with the variable wavenumber

$$k^2(\rho, \theta) = \frac{4k^2 R^2}{[1 + \rho^2 + 2\rho \cos \theta]^2}, \quad (4.23)$$

where the constant k is the physical wavenumber introduced in (4.17).

Thus, we have reduced the original problem (4.17-4.18) formulated on the disk Ω to the Helmholtz equation with variable wavenumber (4.23) to be solved on the semi-plane $0 \leq \theta \leq \pi$. In the framework of Section 4.1.1, this semi-plane can be interpreted as a wedge with the angle

$\omega = \pi$, and the location of singularity in the solution of equation (4.3), (4.23) will be $\rho = 0$. Accordingly, boundary condition (4.22) on the ξ -axis is reformulated as two conditions on two sides of the wedge:

$$\begin{aligned} u|_{\theta=0} &= f(\rho), \\ u|_{\theta=\pi} &= f(-\rho). \end{aligned} \quad (4.24)$$

Next, we express the variable wavenumber (4.23) in the form (4.6). To do so, we expand the right-hand side of (4.23) as a Taylor series with respect to ρ and treat θ as a parameter:

$$k^2(\rho, \theta) = \underbrace{4(kR)^2}_{\equiv k_0(\theta)} + \underbrace{(-16(kR)^2 \sin \theta)}_{\equiv k_1(\theta)} \rho + \underbrace{8(kR)^2(6\sin^2 \theta - 1)}_{\equiv k_2(\theta)} \rho^2 + \dots \quad (4.25)$$

and apply the approach of Section 4.1.1 to the plane wedge with the angle $\omega = \pi$. Having obtained the desired number of terms in the asymptotic expansion of the solution in polar coordinates (ρ, θ) , we transform them into the original (x, y) or (r, φ) coordinates using the inverse of the conformal mapping (4.19). The relationship between (ρ, θ) and (x, y) under mapping (4.19) is

$$\rho(x, y) = \sqrt{\frac{(x - R)^2 + y^2}{(x + R)^2 + y^2}}, \quad (4.26)$$

$$\theta(x, y) = \begin{cases} \arctan \frac{R^2 - x^2 - y^2}{2Ry}, & y > 0, \\ \pi + \arctan \frac{R^2 - x^2 - y^2}{2Ry}, & y < 0. \end{cases} \quad (4.27)$$

As we see from (4.26), $\rho(x, y) \rightarrow 0$ as $x \rightarrow R$ and $y \rightarrow 0$. This is expected since the point $(R, 0)$ is mapped into the origin of the (ξ, η) frame, where $\rho(\xi, \eta) = \sqrt{\xi^2 + \eta^2} \rightarrow 0$ as $\xi, \eta \rightarrow 0$. On the other hand, when $x \rightarrow -R$ and $y \rightarrow 0$, $\rho(x, y)$ becomes unbounded since the point $(-R, 0)$ corresponds to $\pm\infty$ on the real axis ξ . Therefore, expansion (4.7), which is designed to approximate the solution near $(R, 0)$, becomes unbounded at the opposite point $(-R, 0)$ and cannot be used in the vicinity of this point even formally. The latter circumstance is important from the implementation viewpoint (see the discussion in the description of Test 1 in Section 4.1.3).

Later, we will also need to consider the behavior of several functions of ρ and θ in the vicinity of the point $(R, 0)$. The corresponding analysis is straightforward in the (ξ, η) coordinates at the origin $(0, 0)$. Since the mapping (4.19) is analytic near this point, the analytical properties of a given function established in the (ξ, η) coordinates will remain the same in the (x, y) or (r, φ) coordinates. Omitting the simple calculations, we present the following results:

- The function $\theta(x, y)$ of (4.27) is undefined at the point $(R, 0)$ and its first partial deriva-

tives are unbounded there.

- Any combination of $\rho^j \theta$, where j is a non-negative integer, has unbounded derivatives at $(R, 0)$ starting from order $(j + 1)$.
- Products of the kind $\rho^j \sin j\theta$, $\rho^j \cos j\theta$, $\rho^j \sin((j - 2)\theta)$, $\rho^j \cos((j - 2)\theta)$ are regular with all of their derivatives at $(R, 0)$.
- The j -th derivative of the term $\rho^j \ln \rho$ is discontinuous at $(R, 0)$.

Finally, we emphasize that a small region near the vertex of any curvilinear wedge can be conformally mapped onto a small region near the vertex of a straight angle. Therefore, the same approach presented here for the disk will also apply to more complicated geometries.

4.1.3 Test problems

In this section, we introduce six test problems with discontinuous boundary conditions set on the circle of radius R . The first three tests correspond to the Dirichlet problem with a jump discontinuity in (i) the boundary data itself, (ii) the first tangential derivative, and (iii) the second tangential derivative. These Dirichlet problems share the same geometry, location of singular points, and derivative order of singularities studied in the numerical experiments of Section 3.3.11, but differ in the precise boundary condition used - in particular, the loss of the design convergence rate was observed until the boundary data was twice differentiable. The remaining three tests correspond to the Neumann problem with singularities in the boundary data and their first and second tangential derivatives in the same sequence. We discuss Test 1 in greater detail in this section as an example, with the full description of the other test problems presented in Appendix B.

Test 1 We consider piecewise constant Dirichlet data for equation (4.3) with $k^2(\rho, \theta)$ given by (4.23):

$$u|_{r=R} = \begin{cases} 1, & 0 < \varphi < \pi, \\ 0, & \pi < \varphi < 2\pi. \end{cases} \quad (4.28)$$

The Dirichlet data (4.28) undergo a unit jump at $\varphi = 0$ and at $\varphi = \pi$. After the transformation (4.19), boundary condition (4.24) for the piecewise constant data (4.28) reads:

$$\begin{aligned} u|_{\theta=0} &= 1, \\ u|_{\theta=\pi} &= 0. \end{aligned} \quad (4.29)$$

Next, we solve equations (4.8-4.9) consecutively for the first five pairs of coefficients $A_j(\theta)$, $B_j(\theta)$. According to (4.29), we take $f_0 = 1$, $h_0 = 0$, and $f_j = h_j = 0$ for $j = 1, 2, 3, 4$ in the

expansion (4.5), while k_j are given by (4.25). The solution we arrive at is the following:

$$\begin{aligned} A_0(\theta) &= A_1(\theta) = A_3(\theta) \equiv 0, \\ A_2(\theta) &= -\frac{k^2 R^2}{\pi} \sin 2\theta, \\ A_4(\theta) &= \frac{k^2 R^2}{3\pi} [(3 - k^2 R^2/4) \sin 4\theta + k^2 R^2 \sin 2\theta], \end{aligned} \quad (4.30)$$

$$\begin{aligned} B_0(\theta) &= 1 - \frac{\theta}{\pi}, \quad B_1(\theta) = C_1^{(B)} \sin \theta, \\ B_2(\theta) &= C_2^{(B)} \sin 2\theta + k^2 R^2 \left(1 - \frac{\theta}{\pi}\right) (\cos 2\theta - 1), \\ B_3(\theta) &= C_3^{(B)} \sin 3\theta - C_1^{(B)} \frac{1}{2} k^2 R^2 \sin \theta + \frac{2k^2 R^2}{\pi} \left(\frac{1}{2} \sin 2\theta + 1 - \frac{\theta}{\pi}\right) \sin \theta, \\ B_4(\theta) &= C_4^{(B)} \sin 4\theta - C_2^{(B)} \frac{1}{3} k^2 R^2 \sin 2\theta + C_1^{(B)} \frac{4}{3} k^2 R^2 \sin^4 \theta \\ &\quad - \frac{k^2 R^2}{6\pi} \left[(\theta - \pi) (\cos 2\theta - 1) ((k^2 R^2 - 12) \cos 2\theta - k^2 R^2) \right. \\ &\quad \left. - \frac{1}{24} \sin 2\theta (3 \cos 2\theta (k^2 R^2 - 12) - 16k^2 R^2 - 96) \right]. \end{aligned} \quad (4.31)$$

For convenience, we introduce a shorthand notation for the truncated asymptotic expansion at the point $(R, 0)$ [cf. formula (4.7)]:

$$u^{(R,0)} = \sum_{j=0}^4 u_j^{(R,0)}, \quad (4.32)$$

where

$$u_j^{(R,0)} = \rho^j (A_j(\theta) \ln \rho + B_j(\theta)). \quad (4.33)$$

Again, in practice, ρ and θ in formulae (4.30-4.33) should be regarded as functions of the coordinates (x, y) or (r, φ) on the disk \mathbb{D} according to the transformations (4.26-4.27).

Each subsequent term in the sum (4.32) is more regular than the previous term, i.e., a singularity appears in the derivatives of (at least) one order higher than that for the previous term. Indeed, for each individual term (4.33) we can write with the help of (4.30-4.31) and the considerations on regularity presented (in the form of a bulleted list) at the end of Section 4.1.2:

$$u_j^{(R,0)} = \dots \underbrace{\rho^j \ln \rho}_{\substack{j\text{-th derivative} \\ \text{is discontinuous}}} + \dots \underbrace{\rho^j \theta}_{\substack{(j+1)\text{-th derivatives} \\ \text{are discontinuous}}} + \dots \quad (4.34)$$

In formula (4.34), the dots substitute for the unimportant multiplicative constants, as well as for the regular terms. We thus see that the lowest order discontinuous derivative is the j -th derivative that appears in the logarithmic term.

At this point, we can justify why we have truncated expansion (4.7) after $j = 4$, i.e., why we took exactly five terms in expansion (4.32). According to (4.34), with five terms used in (4.32) the difference $v = u - u^{(R,0)} = \sum_{j \geq 5} \rho^j (A_j(\theta) \ln \rho + B_j(\theta))$ will be free of the first five low-order terms that may exhibit a singularity at the point $(R, 0)$, and its expansion will start with the terms $\propto \rho^5$, thereby guaranteeing the continuity of all derivatives up to the fourth order.

Moreover, a direct calculation suggests that the terms $u_j^{(R,0)}$ of (4.33) with the coefficients (4.30-4.31) do not satisfy the Helmholtz equation. Therefore the right-hand side of the regularized problem (4.2) is nonzero in this case (which is also to be expected in general):

$$f = -\mathbf{L}u^{(R,0)} = \mathbf{L}(u - u^{(R,0)}) = \mathbf{L}v. \quad (4.35)$$

In (4.35), we have taken into account that $\mathbf{L}u = 0$ for the exact solution u . The degree of regularity of f given by (4.35) is important from the standpoint of using the compact scheme (2.33), because the five-node stencil applied to f on the right-hand side of (2.33) renders a central difference approximation of the second derivatives of f , see [27, 30, 24]. Those derivatives are guaranteed to be continuous if expansion (4.32) contains at least five terms.

To summarize, we differentiate $v = u - u^{(R,0)}$ four times altogether, with every differentiation reducing the degree of regularity by one. Therefore, to maintain the continuity and boundedness of all quantities employed by our numerical algorithm we need an expansion that is four times continuously differentiable. This is facilitated by taking at least five terms in (4.32).

The boundary data for the regularized problem (4.2) are given by $u - u^{(R,0)}|_{r=R}$. For the current test case, the asymptotic solution $u^{(R,0)}$ takes a very simple form on the circle $r = R$. Indeed, the function $\theta(x, y)$ of (4.27) is equal either to zero (for $y > 0$) or to π (for $y < 0$) at $r = R$. Therefore, according to (4.30)–(4.31), $u^{(R,0)}$ equals one on the upper semi-circle and equals zero on the lower semi-circle. Taking into account the boundary conditions (4.28), we conclude that the Dirichlet boundary data for the regularized problem are zero on the entire circle $r = R$. In general (see Appendix B), the regularity of the boundary data for the regularized problem matches that of the solution to the regularized problem.

It is also to be noted that the expansion of the truncation error for scheme (2.33) starts with sixth-order derivatives of the solution, see [27, 30, 24]. To maintain their boundedness, we would formally need to include additional terms into the sum (4.32) beyond $j = 4$; however, as demonstrated by our numerical experiments in Section 4.3, taking five terms proves sufficient for restoring the design fourth-order convergence rate of the method of difference potentials if

applied to the regularized problem (4.2) or, equivalently, (4.1).

Finally, we note that the asymptotic expansion $u^{(-R,0)}$ at the opposite singular point $(-R, 0)$ can be obtained with no additional effort due to the symmetry of the original problem with boundary data (4.28), and this is accomplished by an even reflection about the y axis: i.e., $x \mapsto -x$, $y \mapsto y$ in the expressions (4.26-4.27). This symmetric configuration has been chosen for simplicity; it presents no loss of generality. To suppress the singularities of the solution at both singular points, $(R, 0)$ and $(-R, 0)$, one should use the sum $u^{(R,0)} + u^{(-R,0)}$ of the corresponding asymptotic expansions.

Implementation notes First of all, the coefficients $B_j(\theta)$ given by (4.31) include some undetermined constants $C_j^{(B)}$ (see equation (4.12) and the discussion that follows). From the standpoint of theory, this uncertainty presents no problem. In practice, however, we must decide how to handle it.

We again refer to the bulleted list at the end of Section 4.1.2 and see that all the terms in the expressions $B_j(\theta)\rho^j$, $j = 0, \dots, 4$, except those that contain θ as a factor are regular at $(R, 0)$ with all their derivatives. We also recall that the sole purpose of using the truncated asymptotic expansion (4.32) is to subtract it from the exact solution and thereby remove the near-boundary singularity or, more precisely, reduce it to a level beyond which the finite difference scheme becomes insensitive to higher-order singular terms. Therefore, keeping the aforementioned regular terms in or dropping them from (4.32) will make no difference as far as achieving our key goal,⁵ which is restoring the design convergence rate of the numerical method. As such, we set the undetermined coefficients $C_j^{(B)}$ to zero for convenience. Moreover, we could have omitted all other terms in the expressions for $B_j(\theta)$ except those containing θ because those terms are regular as well. However, we have chosen to keep them in our tests to demonstrate the fact that having or not having an additional regular component in the expansion will not affect the numerical performance in any way. Indeed, the only essential requirement of the regularized problem (4.1) or (4.2) is that its solution must be sufficiently smooth so as to re-enable the design high order of accuracy of the scheme. If this requirement is met, then the regularized problem can be solved numerically, and the solution to the original problem can subsequently be restored by adding back the previously subtracted singular part.

Next, to compute the right-hand side for the regularized problem, one applies the Helmholtz operator to the sum of the asymptotic expansions, $u^{(R,0)} + u^{(-R,0)}$. Hence, the latter must be known everywhere on the disk Ω . However, as has been mentioned, the function $\rho(x, y)$ given by (4.26) is unbounded at the opposite point $(-R, 0)$, and, consequently, the function $u^{(R,0)}$ is not defined there. Likewise, the function $u^{(-R,0)}$ is not defined at the point $(R, 0)$. Therefore,

⁵The same applies not only to the genuine regular terms but also to the terms that have sufficiently many continuous derivatives.

to achieve the desired regularization, instead of subtracting $(u^{(R,0)} + u^{(-R,0)})$ from the exact solution u , we will first modify $u^{(R,0)}$ and $u^{(-R,0)}$ in a particular way and then subtract:

$$v = u - \left(\mu^{(R,0)} u^{(R,0)} + \mu^{(-R,0)} u^{(-R,0)} \right). \quad (4.36)$$

The multipliers $\mu^{(R,0)}$ and $\mu^{(-R,0)}$ in equation (4.36) are smooth functions equal to unity on some neighborhood of $(R, 0)$ and $(-R, 0)$, receptively. Further away from $(\pm R, 0)$, those multipliers gradually decay to zero. An example of a suitable $\mu^{(R,0)}$ on the unit disk is shown in Figure 4.2. The function $\mu^{(-R,0)}$ is an even reflection of $\mu^{(R,0)}$ around the y axis.

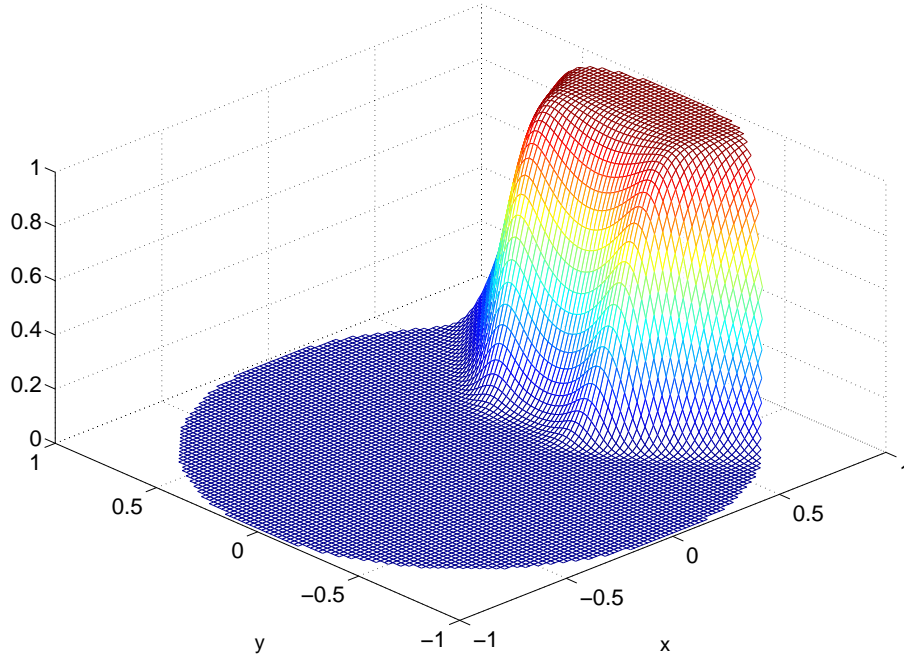


Figure 4.2: Function (4.37) on the disk $R = 1$. The parameters are $r_1 = 0.1$, $r_2 = 0.9$, $\varphi_1 = \pi/12$, $\varphi_2 = \pi/3$

What Figure 4.2 actually represents is the function defined in polar coordinates:

$$\mu^{(R,0)}(r, \varphi) = \mu_r(r) \mu_\varphi(\varphi), \quad (4.37)$$

where

$$\mu_r(r) \equiv P_6 \left(\frac{r - r_1}{r_2 - r_1} \right), \quad (4.38)$$

$$\mu_\varphi(\varphi) \equiv \begin{cases} 1 - P_6 \left(\frac{\varphi - \varphi_1}{\varphi_2 - \varphi_1} \right), & 0 < \varphi < \pi, \\ 1 - P_6 \left(-\frac{\varphi + \varphi_1}{\varphi_2 - \varphi_1} \right), & -\pi < \varphi < 0, \end{cases} \quad (4.39)$$

and

$$P_6(x) = \begin{cases} 0, & x < 0, \\ x^7 (924x^6 - 6006x^5 + 16380x^4 - 24024x^3 + 20020x^2 - 9009x + 1716), & 0 < x < 1, \\ 1, & x > 1. \end{cases} \quad (4.40)$$

The univariate function $P_6(x)$ grows smoothly from zero to one on the interval $[0, 1]$. Its first six derivatives are continuous at both endpoints, $x = 0$ and $x = 1$, where they are all equal to zero, while the seventh derivative undergoes a jump. In (4.38), r_1 and r_2 denote some positive numbers such that $r_1 < r_2 < R$. Therefore, the function $\mu_r(r)$ of (4.38) smoothly increases from zero to one in the radial direction on the annulus $\{r_1 \leq r \leq r_2\}$, i.e., strictly inside the disk $\Omega = \{0 \leq r \leq R\}$. Similarly, the angles φ_1 and φ_2 are chosen so that $0 < \varphi_1 < \varphi_2 < \pi$. Then, the function $\mu_\varphi(\varphi)$ of (4.39) is equal to one in the sector $|\varphi| \leq \varphi_1$, symmetrically decays to zero for $\varphi_1 < |\varphi| < \varphi_2$, and vanishes for $|\varphi| \geq \varphi_2$. Altogether, this guarantees the desired behavior of $\mu^{(R,0)}$ of (4.37).

We emphasize that the multipliers $\mu^{(\pm R,0)}$ introduced in (4.36) will affect both the right-hand side, see formula (4.35), and the boundary data of the regularized problem (4.2). The resulting final formulation of the regularized problem that is solved numerically by the method of difference potentials is presented in Section 4.2.1.

Additional tests The remaining five test cases are described in detail in Appendix B. Out of the five, two are Dirichlet problems with near-boundary singularities of decreasing strength — one is due to a jump discontinuity in the first derivative of the data function and the other is due to a jump discontinuity in the second derivative of the data function. The remaining three are Neumann problems. They also have near-boundary singularities of decreasing strength, due to a jump in the data function itself, its first derivative, and its second derivative, respectively. In Appendix B, we present the boundary conditions for the remaining five test problems (that replace (4.28)) and provide the coefficients of their asymptotic expansions. In this section, we make a few general comments:

- For all the tests, we use symmetry with respect to the y axis to obtain the expansion

$$u^{(-R,0)}.$$

- The undetermined coefficients $C_j^{(B)}$ are set to zero in all tests.
- The multipliers $\mu^{(\pm R,0)}$, see (4.37), are used in all the tests to enable the independent treatment of each individual singular point.
- The degree of smoothness and the number of terms in the expansion are determined as follows. In the Dirichlet tests with the discontinuous first and second derivatives the non-trivial (i.e., non-constant) terms in the asymptotic expansion begin with the functions $u_1^{(\pm R,0)}$ and $u_2^{(\pm R,0)}$, respectively. Hence, the solution in these cases is smoother than that of Test 1, which is natural because the boundary data are smoother and, intuitively, should require less regularization (see the numerical experiments of Section 3.3.11). Nevertheless, even for these smoother settings we should truncate the expansion exactly at the same level (at $u_4^{(\pm R,0)}$ term) as for Test 1 to maintain the desired smoothness of the right-hand side.
- The asymptotic expansion for the Neumann problems takes the form (4.16), with the respective exponents greater by one than those in the Dirichlet expansion (4.7). Therefore, we can truncate the Neumann expansion at one term earlier (at $j = 3$), which proves sufficient for the regularized difference v to have four continuous derivatives.

4.2 Solution by difference potentials in the presence of singularities

4.2.1 Regularized problem

Assume now that we have a boundary value problem (3.1) on the unit disk Ω with boundary condition (3.1b) that is not smooth at the points $\varphi = 0$ and $\varphi = \pi$ on the unit circle Γ . We then build the appropriate singular functions $\{u_j^{(R,0)}\}_{j=1}^M$ via the methods of Section 4.1 (for all the examples of Section 4.1.3 and Appendix B, it proves sufficient to take $M = 4$) and formulate the regularized BVP (4.1) by subtracting the singular functions multiplied by the cutting function $\mu^{(R,0)}$ from the solution u in the original BVP (3.1). We take advantage of the symmetry of the problem to address the opposite singular point $(-R, 0)$ with no additional effort. Collecting the known terms on the right-hand side, we arrive at the regularized BVP, for which we have

previously used the generic notation (4.1) [cf. equations (4.2), (4.35), and (4.36)]:

$$\mathbf{L}u = - \sum_{j=1}^M \mathbf{L}(\mu^{(R,0)} u_j^{(R,0)}) - \sum_{j=1}^M \mathbf{L}(\mu^{(-R,0)} u_j^{(-R,0)}) \stackrel{\text{def}}{=} f, \quad x \in \Omega, \quad (4.41a)$$

$$\mathbf{l}_\Gamma u = \phi_\Gamma - \sum_{j=1}^M \mathbf{l}_\Gamma(\mu^{(R,0)} u_j^{(R,0)}) - \sum_{j=1}^M \mathbf{l}_\Gamma(\mu^{(-R,0)} u_j^{(-R,0)}) \stackrel{\text{def}}{=} \psi_\Gamma. \quad (4.41b)$$

In formula (4.41b), we allow the boundary operator \mathbf{l}_Γ to specify either a Dirichlet or a Neumann condition. By design, both the right-hand side f of the regularized BVP (4.41a) and the regularized boundary data ψ_Γ of (4.41b) are sufficiently smooth on their respective domains, with at least, respectively, two and four continuous derivatives at the singular points $\varphi = 0$ and $\varphi = \pi$ on the boundary Γ .

The method of difference potentials, as described in Chapter 3, can now be applied to the regularized problem (4.41) without degradation of the design fourth-order convergence rate of the overall scheme. Let u_R denote the “regular” solution, i.e., the solution to the regularized problem (4.41) by the method of difference potentials, which, for a particular grid \mathbb{N} on the square auxiliary domain, has its values defined on the nodes \mathbb{N}^+ (recall that this is the set of interior nodes of Ω plus a “fringe”, see Figure 3.1). Then the numerical solution u to the original problem (3.1) (which we have taken to be homogeneous with discontinuous BCs) at the nodes \mathbb{N}^+ is given by simply adding back the singular expressions calculated at these nodes:

$$u|_{\mathbb{N}^+} = u_R + \mu^{(R,0)}|_{\mathbb{N}^+} \sum_{j=1}^M u_j^{(R,0)}|_{\mathbb{N}^+} + \mu^{(-R,0)}|_{\mathbb{N}^+} \sum_{j=1}^M u_j^{(-R,0)}|_{\mathbb{N}^+}. \quad (4.42)$$

Therefore, the algorithm for the method of difference potentials presented in Chapter 3 does not change at all. What changes, rather, is the problem which we solve by difference potentials. In other words, we rephrase the original singular problem (3.1) as the regularized problem (4.1) which takes the specific form (4.41), apply the method of difference potentials to the regularized problem (4.41), and then add back the singular terms as in (4.42).

4.2.2 Solution of multiple problems at low cost

In the numerical simulations of Section 4.3, all of the problems have singularities at the same locations, $\varphi = 0, \pi$, on the unit circle Γ . Hence, they differ only by the right-hand side of the regularized equation (4.41a) and the boundary data ψ_Γ in (4.41b), which result from the singular functions specific to a given problem, see Section 4.1.3 and Appendix B. As discussed in Section 3.3.1, the similarity of these problems means that after solving one such problem the computational complexity of the subsequent problems with different right-hand sides is

substantially reduced, and this was demonstrated and analyzed for problems with multiple right-hand sides in the simulations of Section 3.3.8. In fact, the basis that we use for the following test problems is exactly the one used in Section 3.3.11, and thus the expensive step of computing \mathbf{Q}_H need not be repeated here. Consequently, having already computed \mathbf{Q}_H ahead of time, the following problems are solved at the nominal cost described in Section 3.3.8. Moreover, we emphasize that the change of the type of the boundary condition from Dirichlet to Neumann does not incur any further computational cost either.

4.3 Numerical simulations

We will be solving the homogeneous Helmholtz equation (3.1a) subject to boundary condition (3.1b) on a disk Ω of radius 1, centered at the origin. For the purpose of using the method of difference potentials, Ω is embedded in an auxiliary domain that is a square of side length 2.2, also centered at the origin, see Section 3.1.1. We form series expansions of the boundary data using Chebyshev basis functions as described in Section 3.2. The originally posed problem is modified using the method of singularity subtraction outlined in Section 4.1, with particular singular functions for each problem derived in Section 4.1.3 and Appendix B. This results in an inhomogeneous problem (4.1) or (4.41) which no longer has a discontinuity in the boundary condition and whose right-hand side is sufficiently smooth. The solution of problem (4.41) is therefore expected to possess sufficient regularity so that the method of difference potentials will yield the numerical solution at the design rate of grid convergence for the scheme (2.33), which is fourth order. After computing the numerical solution to this regularized problem, we add back the singular functions to the numerical solution in order to obtain an approximation to solution of the original singular problem (3.1).

4.3.1 Parameters of the computational setting

Scheme and errors In all of the following test problems, the calculations are conducted using the fourth-order accurate compact finite difference scheme (2.33) supplemented by the Sommerfeld-type boundary conditions (2.50a-2.50b) at the left and right edges of the auxiliary square and a Dirichlet condition (2.44a) at its top and bottom edges. These computations are carried out on a series of Cartesian grids containing 64, 128, 256, 512, 1024, and 2048 cells uniformly spaced in each direction, with each grid being nested within the previous. We do not suppose that the exact solution to each problem is known, and thus we use this nesting of the grids to compute the error in the “Cauchy sense,” i.e., the maximum absolute value of the difference between the two numerical solutions on a pair of consecutive grids, with this difference evaluated at the nodes of the coarser grid. The convention which we adopt in Tables 4.1-4.7 is

that the coarser grid involved in the computation of the error is shown, and the convergence rates shown are computed as the binary logarithm of the ratios of successive errors. For the first test, we have also found the exact solution explicitly (Section 4.3.2), which enables a direct study of the grid convergence as an additional validation of our method.

Geometry and basis functions The circular boundary Γ is partitioned into two arcs which meet, by design, at the discontinuity locations of the boundary data, so that $\Gamma_1 = \{r = 1, 0 < \varphi < \pi\}$, $\Gamma_2 = \{r = 1, \pi < \varphi < 2\pi\}$. Thus, the trace of the solution along each arc of the circle is independently represented by its own set of Chebyshev basis functions.⁶ We implement the Chebyshev basis functions on an extended interval $[-1-\varepsilon, 1+\varepsilon]$ in order to avoid numerical difficulties that arise when computing the derivatives of the Chebyshev functions near the endpoints of the interval $[-1, 1]$. In all of the simulations, this parameter is chosen to be $\varepsilon = 0.001$ (for details on the choice of this parameter, see Section 3.3).

The number of basis functions N used to expand the boundary data according to formula (3.33) is chosen specifically for each grid and each problem⁷. On one hand, the number of basis functions must approximate the boundary data with accuracy that matches or exceeds the accuracy of the finite difference scheme (2.52). On the other hand, having too many basis functions on a given grid will result in a loss of accuracy. The reason is that each Chebyshev basis function is more oscillatory than the previous one. Eventually, for a particular grid, a basis function is reached whose oscillations are finer than the grid size, and hence all subsequent basis functions become essentially indistinguishable on the grid. This artificial loss of accuracy is alleviated when moving to finer grids, which may give the false impression of an unusually high convergence rate (see Section 3.3.6). We have observed by trial-and-error that the threshold for the loss of accuracy due to having too many basis functions on the coarsest grid, 64×64 , is around $N = 45$ basis functions for this problem. Rather than choosing an arbitrary tolerance σ for the truncation of the basis expansion of the boundary data ahead of time, as done for the simulations of Section 3.3.2, we present the following alternative strategy of determining the appropriate number of basis functions in each case:

1. We run the simulation first on the coarsest grids, 64×64 and 128×128 , using 45 basis functions on each. By saving the matrix \mathbf{Q}_H from each test, we can very cheaply reduce

⁶These bases need not have the same dimensions: neither the bases on each arc or even the bases for the Dirichlet or Neumann portions of the data need to match. That is, we could instead assign altogether 4 different numbers of basis functions, one for each of the two arcs and one for each component of the Cauchy data on each arc. It is only for convenience that we use the same number of basis functions N on each arc and each component of the data.

⁷Each boundary segment is of equal length in this case and we have no other reason, such as changes in the geometry, to expect to need more basis functions on one or the other. For this reason, we choose to have the same N as the number of basis functions on each respective segment, unlike the discussions in Section 3.3.6.

the number of basis functions in subsequent simulations by simply eliminating the corresponding columns of \mathbf{Q}_H . If we instead wish to add more basis functions, we may also prevent redundant computation by only computing the additional columns of \mathbf{Q}_H that we desire.

2. We run the simulation again on the 64×64 grid with fewer and fewer basis functions⁸, computing the error for each test by comparing the solution to the one on the 128×128 grid. At first the error decreases because the number of basis functions decreases from $N = 45$, which is too many, to smaller values. Then, it reaches what we will refer to as the “grid error” — that is, the error which is free from both the interference due to oscillations in higher basis functions and from the insufficiently accurate approximation of the boundary data. The design rate of grid convergence can be observed if the number of basis functions falls into this middle range (not too many, but enough for approximating the boundary data), which is different for each grid.
3. Once the grid error has been determined, we compute the truncation error of the Chebyshev expansion of the boundary data for the minimum number of basis functions which yields the grid error. This truncation error can be evaluated by computing the expansion with many more basis functions and then looking at the maximum absolute value of the coefficients beyond the chosen point which yields the grid error.
4. Knowing that the finite difference scheme is fourth order accurate, and that our consecutive grids will each have twice as many nodes in each direction, we expect that the grid error for the 128×128 grid will be smaller than that of the 64×64 grid by a factor of 16. Therefore, we obtain the necessary truncation error for the Chebyshev series on the 128×128 grid by dividing the truncation error of the Chebyshev series obtained on the 64×64 grid by 16, and then determine the number of coefficients required to achieve this truncation error on the finer grid. The resulting number of coefficients should be sufficient to achieve the grid error on the 128×128 grid. According to step 1, we may need to compute only the additional columns of \mathbf{Q}_H that correspond to the basis functions beyond $N = 45$, if necessary.
5. For all subsequent grids, we continue to divide the truncation error obtained in step 4 by another factor of 16 (so that for the grid 256×256 , we are dividing the original truncation error by 16^2 , etc.) and then find the appropriate number of Chebyshev terms corresponding to that grid the same way as before.

⁸This is inexpensive since the projection of each basis function has already been computed. Therefore, we may simply drop off the appropriate columns of the matrix \mathbf{Q}_H .

Remarks Because these grids are coarse, the computational cost of this procedure, in view of the goal of determining a sufficient number of basis functions for finer grids, is comparatively very low. Note that if the Chebyshev expansions of the boundary functions converge too slowly, then the minimum number of basis functions required to represent the boundary data may exceed the capability of the grid to resolve the basis functions. However, this has never been observed in practice and is, in fact, prevented by design. Recall that the convergence of the Chebyshev series depends on the smoothness of the function being expanded: it is noted in Section 3.2.5 that the Chebyshev coefficients decay relative to the degree of smoothness of the function similarly to the Fourier coefficients (see footnote 3, page 44). While the original boundary data in our problems are discontinuous, the boundary data are “smoothed” by the process of singularity subtraction, so that the data on which we perform the expansion have at least 4 continuous derivatives. This has proven to be sufficient for our case, as the results in Tables 4.1-4.7 confirm. Since the number of basis functions on each grid depends on the boundary data of the problem, our particular choices are displayed in Tables 4.1-4.7.

The computer implementation of the algorithm is performed in MATLAB.

4.3.2 Test 1: Discontinuous Dirichlet boundary data

The specified Dirichlet condition is discontinuous,

$$u|_{r=1} = \begin{cases} 1, & 0 < \varphi < \pi, \\ 0, & \pi < \varphi < 2\pi, \end{cases}$$

and the coefficients of the singular terms are given by (4.30-4.31). We demonstrate the convergence of the method in Table 4.1 for the wavenumbers $k = 5$ and $k = 15$.

Table 4.1: Results for Dirichlet boundary data with discontinuity.

Grid	$k = 5$			$k = 15$		
	N	Error	Conv. rate	N	Error	Conv. rate
64×64	30	2.46	-	40	0.73	-
128×128	30	$6.40 \cdot 10^{-4}$	11.91	40	$4.72 \cdot 10^{-2}$	3.96
256×256	45	$2.48 \cdot 10^{-5}$	4.69	50	$3.02 \cdot 10^{-3}$	3.97
512×512	70	$1.58 \cdot 10^{-6}$	3.97	53	$1.91 \cdot 10^{-4}$	3.99
1024×1024	81	$3.60 \cdot 10^{-7}$	2.13	71	$1.19 \cdot 10^{-5}$	4.00

As an additional corroboration of the performance of the method, we present a comparison to the exact solution for this test problem:

$$u(r, \varphi) = \frac{1}{2} \frac{J_0(kr)}{J_0(kR)} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{J_{2n+1}(kr)}{J_{2n+1}(kR)} \frac{\sin(2n+1)\varphi}{2n+1}, \quad (4.43)$$

where J_n are the Bessel functions of the first kind and R is the radius of the circle, which in our case is $R = 1$. In practice, we must make a few modifications in order to evaluate (4.43) with sufficient accuracy. First, we truncate the series expansion at 2000 terms. Next, for terms of the series beyond the 50-th we replace the ratio of Bessel functions by its asymptotic form r^{2n+1} due to a loss of numerical stability in computing higher-order Bessel functions. Therefore, we approximate the solution as follows:

$$u(r, \varphi) = \frac{1}{2} \frac{J_0(kr)}{J_0(kR)} + \frac{2}{\pi} \left[\sum_{n=0}^{50} \frac{J_{2n+1}(kr)}{J_{2n+1}(kR)} \frac{\sin(2n+1)\varphi}{2n+1} + \sum_{n=51}^{2000} r^{2n+1} \frac{\sin(2n+1)\varphi}{2n+1} \right]. \quad (4.44)$$

One final note is that the series (4.43) converges poorly near the boundary of the disk, $R = 1$. To overcome this, we compare the numerical solution to the expansion (4.44) only on the interior 90% of the disk (i.e., on the subset $r < 0.9$). The results of this comparison are summarized in Table 4.2.

Table 4.2: Results for Dirichlet boundary data with discontinuity. The error is computed by comparison to the approximation (4.44) of the exact solution.

Grid	$k = 5$			$k = 15$		
	N	Error	Conv. rate	N	Error	Conv. rate
64×64	30	$2.4 \cdot 10^{-3}$	-	40	0.81	-
128×128	30	$1.39 \cdot 10^{-4}$	4.11	40	$5.58 \cdot 10^{-2}$	3.86
256×256	45	$2.33 \cdot 10^{-6}$	5.90	50	$3.21 \cdot 10^{-3}$	4.12
512×512	70	$1.15 \cdot 10^{-7}$	4.34	53	$2.03 \cdot 10^{-4}$	3.99
1024×1024	81	$8.38 \cdot 10^{-9}$	3.78	71	$1.27 \cdot 10^{-5}$	3.99

4.3.3 Test 2: Continuous Dirichlet boundary data with first derivative discontinuity.

For this problem, the Dirichlet condition is discontinuous in the first derivative,

$$u(\varphi) = \begin{cases} \frac{\pi}{2} - \varphi, & 0 < \varphi < \pi, \\ \varphi - \frac{3\pi}{2}, & \pi < \varphi < 2\pi, \end{cases}$$

and the coefficients of the singular terms are given by (B.3-B.4).

Table 4.3: Results for continuous Dirichlet boundary data with first derivative discontinuity.

	$k = 1$			$k = 5$		
Grid	N	Error	Conv. rate	N	Error	Conv. rate
64×64	40	0.39	-	40	0.80	-
128×128	50	$7.42 \cdot 10^{-5}$	12.36	40	$1.83 \cdot 10^{-4}$	12.09
256×256	66	$5.60 \cdot 10^{-6}$	3.73	50	$7.18 \cdot 10^{-6}$	4.67
512×512	82	$3.16 \cdot 10^{-7}$	4.15	74	$4.42 \cdot 10^{-7}$	4.02
1024×1024	111	$2.10 \cdot 10^{-8}$	3.91	111	$2.63 \cdot 10^{-8}$	4.07

4.3.4 Test 3: Continuous Dirichlet boundary conditions with second derivative discontinuity.

The Dirichlet boundary condition is now discontinuous in the second derivative,

$$u(\varphi) = \begin{cases} \cos \varphi, & 0 < \varphi < \pi \\ \cos 3\varphi, & \pi < \varphi < 2\pi, \end{cases}$$

with the coefficients of the singular terms given by (B.7-B.8).

Table 4.4: Results for continuous Dirichlet boundary data with second derivative discontinuity.

	$k = 1$			$k = 5$		
Grid	N	Error	Conv. rate	N	Error	Conv. rate
64×64	40	0.97	-	40	7.55	-
128×128	50	$4.34 \cdot 10^{-4}$	11.12	40	$1.19 \cdot 10^{-4}$	15.95
256×256	66	$3.82 \cdot 10^{-5}$	3.50	56	$1.48 \cdot 10^{-5}$	3.01
512×512	76	$2.33 \cdot 10^{-6}$	4.03	86	$7.77 \cdot 10^{-7}$	4.25
1024×1024	110	$1.41 \cdot 10^{-7}$	4.05	124	$4.76 \cdot 10^{-8}$	4.03

4.3.5 Test 4: Discontinuous Neumann boundary data

The discontinuous Neumann boundary condition for this test is

$$\frac{\partial u}{\partial n}(\varphi) = \begin{cases} 1, & 0 < \varphi < \pi, \\ 0, & \pi < \varphi < 2\pi, \end{cases}$$

and the coefficients of the singular terms are given by (B.11-B.12).

Table 4.5: Results for Neumann boundary data with discontinuity.

	$k = 1$			$k = 5$		
Grid	N	Error	Conv. rate	N	Error	Conv. rate
64×64	40	0.26	-	40	3.57	-
128×128	50	$7.60 \cdot 10^{-5}$	11.71	50	$1.30 \cdot 10^{-4}$	14.73
256×256	71	$6.03 \cdot 10^{-6}$	3.65	57	$1.51 \cdot 10^{-5}$	3.11
512×512	96	$2.62 \cdot 10^{-7}$	4.52	90	$8.67 \cdot 10^{-7}$	4.12
1024×1024	124	$1.47 \cdot 10^{-8}$	4.15	119	$7.11 \cdot 10^{-8}$	3.61

4.3.6 Test 5: Continuous Neumann boundary data with first derivative discontinuity.

In this test, the Neumann boundary data are discontinuous in the first derivative,

$$\frac{\partial u}{\partial n}(\varphi) = \begin{cases} \frac{\pi}{2} - \varphi, & 0 < \varphi < \pi, \\ \varphi - \frac{3\pi}{2}, & \pi < \varphi < 2\pi, \end{cases}$$

and the coefficients of the singular terms are given by (B.13-B.14).

Table 4.6: Results for continuous Neumann boundary data with first derivative discontinuity.

	$k = 1$			$k = 5$		
Grid	N	Error	Conv. rate	N	Error	Conv. rate
64×64	40	0.55	-	40	0.52	-
128×128	50	$1.78 \cdot 10^{-5}$	14.92	50	$6.57 \cdot 10^{-4}$	9.64
256×256	76	$1.85 \cdot 10^{-6}$	3.26	71	$1.35 \cdot 10^{-5}$	5.61
512×512	94	$1.90 \cdot 10^{-7}$	3.28	96	$2.90 \cdot 10^{-7}$	4.78
1024×1024	128	$8.97 \cdot 10^{-9}$	4.40	122	$2.66 \cdot 10^{-8}$	4.20

4.3.7 Test 6: Continuous Neumann boundary data with second derivative discontinuity.

For the final test, the Neumann data have a jump in the second derivative,

$$\frac{\partial u}{\partial n}(\varphi) = \begin{cases} \cos \varphi, & 0 < \varphi < \pi, \\ \cos 3\varphi, & \pi < \varphi < 2\pi, \end{cases}$$

and the coefficients of the singular terms are shown in (B.15-B.16).

Table 4.7: Results for continuous Neumann boundary data with second derivative discontinuity.

	$k = 1$			$k = 5$		
Grid	N	Error	Conv. rate	N	Error	Conv. rate
64×64	40	0.83	-	40	5.10	-
128×128	40	$5.76 \cdot 10^{-4}$	10.49	40	$6.03 \cdot 10^{-4}$	13.05
256×256	60	$3.23 \cdot 10^{-5}$	4.15	60	$3.47 \cdot 10^{-5}$	4.12
512×512	96	$2.32 \cdot 10^{-6}$	3.80	96	$1.65 \cdot 10^{-6}$	4.39
1024×1024	126	$1.37 \cdot 10^{-7}$	4.08	128	$6.62 \cdot 10^{-8}$	4.64

4.4 Discussion

We have shown in this chapter how to apply the method of difference potentials when computing singular solutions of the Helmholtz equation while preserving high-order accuracy. The key idea is to regularize the original problem by subtracting several leading terms of the asymptotic expansion of the solution near the singularity, which we have demonstrated for the Dirichlet and Neumann boundary conditions, and to then approximate numerically only the remaining sufficiently smooth part of the solution. Additionally, we have implemented a local conformal mapping which reduces the problem of finding the coefficients of the asymptotic expansion to the simple case of a wedge for curvilinear boundaries with accuracy sufficient to retain the high order of the overall method. As a result, the finite difference scheme maintains its consistency for the regularized problem, and, as the computations in Section 4.3 demonstrate, the overall numerical method converges with the design rate. In doing so, the method of difference potentials has also permitted us to handle non-conforming curvilinear boundaries on regular structured grids with no deterioration of accuracy and enables the solution of a series of problems with various boundary conditions at a low computational cost per problem (Sections 3.3.1, 4.2.2). Tables 4.1-4.7 demonstrate that high order accuracy is maintained for a variety of wavenumbers and boundary conditions with singularities of different strengths.

It may be noted, however, that the number of basis functions required in the simulations of this chapter are considerably higher than in the simulations of Chapter 3, with 128 basis functions on each arc of the Chebyshev basis in Section 4.3.7 for a total of $4(128) = 512$ basis functions. The reason for this is because even the regularized problem contains a weak singularity at the boundary points, resulting in slower convergence of the Chebyshev series. This is only one of many factors that may affect the total number of basis functions, and these are explored in some detail in Chapter 5.

Chapter 5

The Pollution Effect and Efficiency of the Algorithm

The most expensive step of the difference potentials algorithm presented in Chapter 3 is that of applying the difference projection to each basis function of the series representation of the boundary data. For each additional basis function, applying the projection entails solving the AP (2.51) by finite differences with a new right-hand side g . The central dilemma in choosing the minimally sufficient number of basis functions on a particular grid is the lack of an error estimate to provide guidance on the tolerance for the basis expansion of the boundary data. There are additional factors that determine the actual number of basis functions needed, but in fact it is the truncation tolerance σ of the basis expansion that we must be concerned with rather than the specific number of basis functions required to meet that accuracy for a particular problem.

For the Helmholtz equation, the dispersion error, or pollution effect, is known to behave proportional to h^4k^5 asymptotically as the step size h tends to zero and the wavenumber k increases [1, 2, 3]. The behavior of the coefficient of h^4k^5 in the asymptotic expansion of the error is nontrivial and difficult to obtain. Rather than attempt to obtain an error estimate by analyzing the complicated behavior of this constant, we instead assume that the error may be estimated as a coefficient C_P times the leading term, $C_P h^4k^5$, where C_P is not the exact constant coefficient of the h^4k^5 in the asymptotic expansion of the error but instead accounts for the behavior of the higher order terms as well and thus may vary. Using this simplification¹, we investigate the behavior of C_P for the range of problems we have solved thus far.

¹This differs from our study of the pollution effect in [23], in which we confirm the pollution effect for a fourth order polar FD scheme for the Helmholtz equation. In this work, instead of demonstrating the contribution of the pollution effect to the error we assume that the overall error is dominated by the dispersion error. This provides an *a priori* approximation of the error within a range of related problems once a representative sample of them has been solved.

This chapter is divided as follows. In Section 5.1, we experimentally evaluate the pollution coefficient C_P for the variety of simulations conducted in Chapters 3-4 and provide an additional numerical study verifying that this coefficient does not depend significantly on the wavenumber k . Section 5.2 provides a summary analysis of the different procedures employed thus far for choosing the number of basis functions, describes an alternative procedure based on estimation of the pollution coefficient C_P , and demonstrates the relationship between the truncation error σ of the basis expansion of the boundary data and the overall error of the scheme by a numerical example. Finally, in Section 5.3, we implement the alternative strategy of Section 5.2 using the average $\overline{C_P}$ of the experimentally evaluated pollution coefficients C_P from Section 5.1 to determine the number of basis functions for a new problem.

5.1 Numerical investigation of the Pollution Effect

We begin the investigation of the pollution effect with a simple calculation which estimates the pollution coefficient C_P from the previously computed numerical examples of Sections 3.3.3-3.3.8 and Sections 4.3.2-4.3.7. Because of the asymptotic dependence of the pollution effect on not only the step size h but also the wavenumber k , we have omitted the cases from the aforementioned sections with $k = 1$. This set of problems is diverse in several ways, involving different wavenumbers, boundary conditions, singularities, and inhomogeneities for the BVP (3.1). The calculation of the pollution coefficient C_P is performed by dividing the experimentally obtained error for each problem by $h^4 k^5$, and the results are displayed in Table 5.1, with each C_P being computed from the finest available mesh of the cited section.

We observe some similarity in the pollution coefficients of problems solved by the method of difference potentials regardless of the boundary conditions on the circle, the wavenumber, or the inhomogeneous right-hand side of the Helmholtz equation. The coefficient C_P for all of the problems from Chapters 3-4 is consistently between the orders of 10^{-1} and 10^{-3} .

Commonalities of the problems The numerical examples of Sections 3.3.3-3.3.8 and Sections 4.3.2-4.3.7 all involve the constant-coefficient Helmholtz equation on the same domain, the unit disk centered at the origin, and the auxiliary problems being solved by the finite difference scheme are therefore also closely related. Not only are the APs posed on square domains of equal size, but notice that when the boundary conditions on the domain of interest Ω change that the boundary conditions on the encapsulating auxiliary domain Ω_0 remain the same, see (2.51). In effect, the APs utilized throughout all of the problems in Chapters 3-4 differ only in two important ways. First, the right-hand side g of the AP (2.51) changes, and this change comes from a combination of the BCs on the domain Ω , the right-hand side f of the interior BVP

Table 5.1: Estimations of the pollution effect coefficient C_P for various simulations.

Section	Step Size	k	Error	C_P
3.3.3	2.2/2048	10	$8.25 \cdot 10^{-9}$	$6.20 \cdot 10^{-2}$
3.3.4, Table 3.3	2.2/2048	10	$3.05 \cdot 10^{-9}$	$2.29 \cdot 10^{-2}$
3.3.4, Table 3.4	2.2/2048	10	$1.29 \cdot 10^{-9}$	$9.69 \cdot 10^{-2}$
3.3.5, Table 3.5	2.2/2048	10	$4.73 \cdot 10^{-9}$	$3.55 \cdot 10^{-2}$
3.3.5, Table 3.6	2.2/2048	10	$3.07 \cdot 10^{-9}$	$2.31 \cdot 10^{-2}$
3.3.8, Table 3.9	2.2/2048	10	$1.21 \cdot 10^{-9}$	$9.09 \cdot 10^{-3}$
3.3.8, Table 3.10	2.2/2048	10	$1.33 \cdot 10^{-9}$	$9.99 \cdot 10^{-3}$
4.3.2, Table 4.2	2.2/1024	5	$8.38 \cdot 10^{-9}$	$1.26 \cdot 10^{-1}$
4.3.2, Table 4.2	2.2/1024	15	$1.27 \cdot 10^{-5}$	$7.85 \cdot 10^{-1}$
4.3.3, Table 4.3	2.2/1024	5	$2.63 \cdot 10^{-8}$	$3.95 \cdot 10^{-1}$
4.3.4, Table 4.4	2.2/1024	5	$4.76 \cdot 10^{-8}$	$7.15 \cdot 10^{-1}$
4.3.5, Table 4.5	2.2/1024	5	$7.11 \cdot 10^{-8}$	1.07
4.3.6, Table 4.6	2.2/1024	5	$2.66 \cdot 10^{-8}$	$4.00 \cdot 10^{-1}$
4.3.7, Table 4.7	2.2/1024	5	$6.62 \cdot 10^{-8}$	$9.94 \cdot 10^{-1}$

(3.1), and the basis functions chosen, as these right-hand sides are all defined² by the difference potential of equation (3.2). Second, the differential operator \mathbf{L} of the AP (2.51) changes when the wavenumber k of the Helmholtz equation (3.1a) changes. However, we expect this change to be reflected in the term k^5 of the error estimate $C_P h^4 k^5$.

The data in Table 5.1 indicate that the similarities between the APs utilized in the method of difference potentials may allow us to estimate the pollution coefficient for a range of related problems. While the wavenumber does not change throughout Sections 3.3.3-3.3.8, the boundary conditions and even the inhomogeneous right-hand side of the equation vary without much difference in the pollution coefficient C_P . This is also the case in the simulations of Sections 4.3.2-4.3.7, which also have similar pollution coefficients with differences in the boundary conditions and right-hand sides, and also a few differences in the wavenumber. What lacks the most in variety from Table 5.1 is the wavenumber k involved, and to this end we design the numerical experiment of Section 5.1.1. Our expectation, given that the estimate $C_P h^4 k^5$ of the pollution effect already contains the term k^5 , is that the effect on C_P due to the higher order terms of the asymptotic expansion of the error will be small provided the wavenumber k is sufficiently

²Recall that each of these three aspects of the problem will result in different right-hand sides g for the AP (2.51). If the BVP (3.1) is inhomogeneous, then (3.6) implies that the AP must be solved with a right-hand side g that is influenced by this inhomogeneity, and an additional right-hand side of the AP comes from the term $\mathbf{Q}_I f$ of (3.23). Several more right-hand sides g for the AP result from the projection of the basis functions to form \mathbf{Q}_H in (3.23). One final right-hand side g for the AP is due to the final projection required in (3.6) to reconstruct the solution on the interior domain from the boundary data, which is directly affected by the BC on Ω .

large.

5.1.1 Effect of the wavenumber on the pollution coefficient

We design an experiment to isolate and observe the effect of changing the wavenumber on the pollution coefficient C_P by solving several problems which are identical except that each has a different wavenumber ranging from $k = 3, \dots, 26$. By a simple calculation one may verify that there is little difference between the pollution coefficients on the finest grids (2048×2048) and the 512×512 grids for the same problems analyzed in Table 5.1. Therefore, we assume for this experiment that the wavenumber $k \geq 3$ is sufficiently large and the step size $h = 2.2/512 \approx 0.004$ is sufficiently small so that the error is reasonably approximated as $C_P h^4 k^5$.

Parameters We solve the homogeneous Helmholtz equation (3.1a) on the unit disk centered at the origin so that the boundary curve Γ is the unit circle. We specify a Dirichlet BC (3.1b) by truncating a plane wave test solution $u = e^{ikx}$ to the boundary curve Γ . For each wavenumber $k = 3, \dots, 26$, the error is computed in the infinity norm on the domain Ω by comparison to the exact solution $u = e^{ikx}$, and the pollution coefficient C_P in each case is experimentally evaluated from the resulting errors on a 512×512 grid.

We implement a split Chebyshev basis on the segments $\Gamma_1 = \{r = 1, \theta \in [0, 2\pi/3)\}$, $\Gamma_2 = \{r = 1, \theta \in [2\pi/3, 2\pi)\}$. The number of basis functions was chosen to be $N_1 = 55$ and $N_2 = 78$ on Γ_1 and Γ_2 , respectively, for all tests. Experimentally, we have found that these numbers of basis functions allow for accurate computation across this range of wavenumbers at this grid size.

Note that for each different wavenumber the finite difference operator $\mathbf{L}^{(h)}$ will be different. Therefore, the several back solves (specifically, $2N_1 + 2N_2 = 2(55) + 2(78) = 266$) must be computed to form the matrix \mathbf{Q} separately for each wavenumber k (see Table 3.2) even though the basis does not change. For the particular wavenumber $k = 10$ we may reuse the matrix \mathbf{Q} from the previous simulations of Section 3.3.4, which we have also used in several subsequent examples.

The results are summarized in Table 5.2 and visualized in Figure 5.1 along with the mean and median values of C_P from Table 5.2, which are, respectively, 0.072 and 0.032.

5.1.2 Remarks and discussion

Table 5.1 suggests that the pollution coefficient does not vary dramatically among problems solved by difference potentials which share similarities in the AP on Ω_0 but which may differ greatly in the BCs, source terms, and wavenumbers of the original interior problem on $\Omega \subset \Omega_0$. Due to the lack of variety in the wavenumber for the cases analyzed in Table 5.1, the

Table 5.2: Estimation of the pollution coefficient C_P for $k = 3, \dots, 26$ on a 512×512 grid.

k	$\ u - u_{num}\ _\infty$	C_P	k	$\ u - u_{num}\ _\infty$	C_P
3	$3.14 \cdot 10^{-9}$	0.038	15	$7.77 \cdot 10^{-6}$	0.030
4	$5.84 \cdot 10^{-9}$	0.017	16	$1.34 \cdot 10^{-5}$	0.038
5	$1.77 \cdot 10^{-8}$	0.017	17	$1.87 \cdot 10^{-4}$	0.39
6	$1.94 \cdot 10^{-8}$	0.0073	18	$3.95 \cdot 10^{-5}$	0.061
7	$8.89 \cdot 10^{-7}$	0.16	19	$6.72 \cdot 10^{-5}$	0.080
8	$8.05 \cdot 10^{-8}$	0.0072	20	$2.20 \cdot 10^{-4}$	0.20
9	$1.54 \cdot 10^{-7}$	0.0076	21	$3.43 \cdot 10^{-5}$	0.025
10	$8.29 \cdot 10^{-7}$	0.024	22	$5.92 \cdot 10^{-5}$	0.034
11	$1.51 \cdot 10^{-6}$	0.028	23	$5.81 \cdot 10^{-5}$	0.027
12	$1.09 \cdot 10^{-6}$	0.013	24	$2.46 \cdot 10^{-4}$	0.091
13	$1.54 \cdot 10^{-5}$	0.12	25	$3.43 \cdot 10^{-4}$	0.024
14	$1.45 \cdot 10^{-6}$	0.0079	26	$8.92 \cdot 10^{-4}$	0.024

numerical experiment of Section 5.1.1 was devised to more decidedly confirm that C_P does not vary significantly with the wavenumber, as expected. Even when the boundary conditions, inhomogeneous right-hand sides, and wavenumbers of the cases compared in Tables 5.1-5.2 differed widely, the boundary conditions of the AP (2.51) were unchanged — what did change for the AP in every case were the right-hand sides g . This would indicate that among problems which share such similarities in the AP, one may rely on the same estimate of the pollution coefficient across a broad class of problems with different boundary conditions, right-hand sides, and wavenumbers.

Two potentially influential contributions to the pollution coefficient C_P that we have not explored are related to the geometry of both the domain Ω and the auxiliary domain Ω_0 — further investigation is needed in this regard, and an illustration of this contribution is given at the beginning of Section 5.2. Of greater significance, our present observations position us to comment on problems of a narrower focus in which the geometries involved are the same, and this is the ultimate goal of Section 5.2.

5.2 Choosing the Number of Basis Functions

There are at least two important factors which will determine the required number of basis functions: the geometry of the boundary curve and the expected error of the problem.

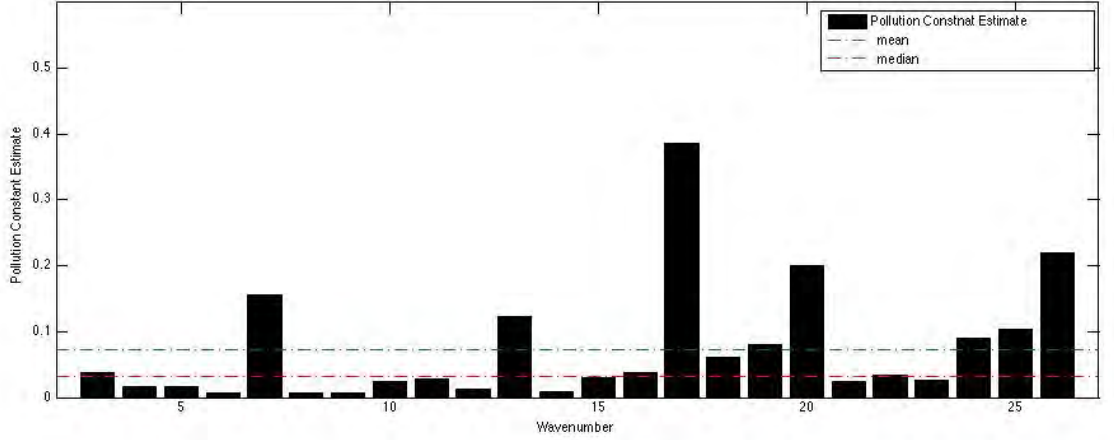


Figure 5.1: Estimated values of the pollution coefficient C_P for $k = 3, \dots, 26$ from the 512×512 grid.

Geometry Firstly, the geometry of the domain will obviously influence the variation of the solution and its normal derivative along the boundary curve. As a simple example, consider a plane wave test solution $u = e^{ikx}$, which we have used in several examples on a circular boundary of radius 1. In this case, the Dirichlet data parameterized by the arc length which we seek to approximate by a series expansion is $u|_{r=1} = e^{ik \cos \theta}$ since $x = \cos \theta$ on the unit circle, where θ coincides with the arc length. Even if the test solution does not change, more complicated geometry of the boundary curve will result in a more complicated expression for the trace of the solution.

One can imagine that a boundary curve of non-constant curvature will inherently imply more variation in the solution and its normal derivative than the circle, and therefore will generally require more basis functions to achieve the same accuracy as on a simpler shape. However, this is something over which we have no control. What we are able to control, rather, is the accuracy σ of the basis representation, and the ideal choice of this tolerance is completely unaffected by the shape of the boundary curve.

Expected error Another major contribution to the number of basis functions required will be the error of the overall scheme, since this will determine the accuracy σ with which we must approximate the boundary data. We expect from the pollution effect that the error will be approximately $C_P h^4 k^5$; thus, the unknown magnitude of the pollution coefficient C_P is a key factor in determining the largest acceptable tolerance σ — i.e., the tolerance which permits the smallest acceptable number of basis functions for the series expansion of the boundary

data and thereby yields the greatest efficiency of the difference potential algorithm detailed in Section 3.2.7.

We may firstly observe that the number of basis functions required will increase as the wavenumber k increases. As remarked upon regarding the geometry, this is due to the simple fact that the boundary data are behaving in a more oscillatory way as the wavenumber k increases, and more basis functions are required in order to resolve that behavior to the same tolerance σ . Even if we adopt a tolerance σ which grows proportional to the expected error, this may still be the case: while the basis expansion is asymptotically expected to converge rapidly for smooth functions, there is no guarantee on the precise point at which this asymptotic behavior begins.

Therefore we desire an estimate of the minimally sufficient number of basis functions, since each additional basis function directly impacts the efficiency of the overall algorithm by adding to the number of back solves required (see Table 3.2); however, as we have endeavored to show, there are problem-specific factors (e.g., the geometry of the domain and wavenumber) that restrict us to quantifying the tolerance σ . In Section 5.2.1 we discuss the strategies that we have employed toward this end in Chapters 3-4 and present another alternative based on the error estimate, and in Section 5.2.2 we provide additional insight into the choice of the tolerance as it relates to the error estimate through a numerical investigation.

5.2.1 Strategies

Strategy 1: fixed predetermined tolerance Even in the most naive approach that we have employed in Chapter 3 before considering how to reduce the total number of basis functions in Sections 3.3.6-3.3.7, the rapid convergence of the basis for smooth functions dictated that the number of basis functions was very small compared to the grid size, meaning that the overall execution time of the algorithm indeed scales proportional to the cost of solving the AP by finite differences. Nevertheless, we showed in Section 3.3.6 that we could reduce the total number of basis functions substantially by relaxing this tolerance a bit after we knew what the error was going to be. In Section 3.3.7 we showed that for problems which utilize the piecewise parameterization of the boundary curve Γ we may also reduce the number of computations by allowing for different numbers of basis functions on separate boundary segments.

The idea of Section 3.3 was to choose the tolerance ahead of time, which we took to be $\sigma = 10^{-10}$. After we had already performed the same computations on all grid sizes with this predetermined tolerance on all grids, in Section 3.3.6 we sought to reduce the number of basis functions on coarser grids. We chose the new tolerance for the finest grid, σ_{2048} , to be equal to the error which had already been obtained on that grid using the original tolerance of $\sigma = 10^{-10}$. Next, we extrapolated backwards to set a tolerance for the coarser grids by multiplying the

tolerance σ_{2048} by the expected factor of 16^3 to obtain the tolerance σ_{1024} for the next coarsest grid, and so on for each subsequent grid.

The choice $\sigma = 10^{-10}$ was made arbitrarily based on our experience (rather than from any kind of analysis) that the error of the FD scheme for these problems would not be smaller than this tolerance even on the finest grid. As long as the tolerance is sufficiently small, the largest contribution to the overall error of the difference potentials algorithm (see Section 3.2.7) is attributable to the FD scheme, which has a design convergence rate of $\mathcal{O}(h^4)$. For greater peace of mind, we could have chosen a tolerance closer to the machine precision, e.g. $\sigma = 10^{-16}$, at the cost of several additional basis functions. Thus it becomes clear that determining the optimal choice of the tolerance σ is precisely the issue in maintaining the optimal efficiency of the overall algorithm.

Strategy 2: *a posteriori* tolerance extrapolation from a coarse grid The method of Section 4.3.1 represents an improvement on the strategy of choosing the tolerance ahead of time. By the same logic employed to reduce the number of basis functions in Section 3.3.6 *a posteriori*, the strategy outlined in Section 4.3.1 is to perform the same process in reverse by obtaining the error on the coarsest grid first and then extrapolating forward to the finer grids. This involves solving the problem on the coarsest grid several times - once with a large number of basis functions, and then several more times, removing one basis function at a time, until we achieved what seems to be the lowest possible error before the overall accuracy of the scheme began to suffer due to insufficient accuracy of the series expansion of the boundary data. Once that point is reached, the magnitude of the last Chebyshev coefficient in the expansion is determined, which provides an estimate for the truncation tolerance σ_{64} of the Chebyshev expansion needed to maintain the accuracy on this coarse grid. From there we extrapolate the truncation tolerance σ_{64} to successively finer grids by dividing by a factor of 16 each time the grid step size h was divided by 2. Though this may appear labor-intensive since we are solving the problem several times, the fact that this is done on a coarse mesh makes it quite an efficient approach if one ultimately desires a mesh that is several times finer. This approach proved to be successful for the simulations of Section 4.3.

One distinction to highlight between these two approaches is that in Strategy 1 we use the error itself as the new tolerance - that is, we assume that the accuracy of the basis expansion of the boundary data needs to be similar to or exceed that of the error on the grid. In Strategy 2, we determine the tolerance empirically by removing basis functions until the point when the series expansion is no longer sufficient. An investigation of the relationship between the truncation tolerance and the error was not presented. Instead, Section 5.2.2 provides a numerical case

³Due to the 4th order convergence rate of the scheme. Since the errors are expected to be $\mathcal{O}(h^4)$, then doubling the step size for each coarser grid is expected to be $\mathcal{O}(2h)^4 = \mathcal{O}(16h^4)$

study in which we compare the empirically determined tolerance for a given problem to the error obtained by the overall scheme.

Strategy 3: determining a tolerance by estimating the pollution coefficient A third approach to setting a suitable tolerance σ would require an *a priori* estimate of the error. As we have remarked, this is difficult in general due to uncertainty in the magnitude of the pollution coefficient C_P . However, we have established in Section 5.1 that the pollution coefficient does not appear to vary greatly for a large class of related problems. Therefore, once a representative problem is solved on a given domain Ω with the AP defined on Ω_0 , the resulting error may be used to compute an estimate of the pollution coefficient C_P that may then be used to compute an *a priori* estimate for future problems on the same domain with the same AP.

In the case that one wishes to solve a variety of problems on a given domain with different wavenumbers, boundary conditions, and inhomogeneous source terms, it may be sufficient to use Strategy 2 for the first few problems and calculate an average $\overline{C_P}$ of the pollution coefficients. This then allows one to estimate the error for subsequent problems as $\overline{C_P}h^4k^5$ regardless of changes in the boundary condition, source term, or wavenumber, and from this error estimate the tolerance σ can be set. The numerical case study of Section 5.2.2 provides evidence that taking σ equal to the overall error of the scheme is sufficient, and this would dictate that we take $\sigma = \overline{C_P}h^4k^5$ when $\overline{C_P}$ is available.

Maximum number of basis functions As a final point of this discussion, we must remember that taking too many basis functions will degrade their linear independence on the discrete boundary γ and lead to a loss of accuracy for coarser grids. Previously, this was observed when the number of basis functions was chosen according to Strategy 1 in Chapter 3, and the goal of Section 3.3.6 was to restore accuracy for the coarser grids by removing several basis functions. The case may arise, however, when the number of basis functions required to achieve the desired accuracy σ chosen according the error estimate in Strategy 3 may exceed the maximum number of basis functions which can be resolved at the discrete boundary γ . Therefore, we will now consider a guideline for setting the maximum number of basis functions on a given segment of the boundary curve.

Without loss of generality, let us consider the segment $\Gamma_1 = \Gamma|_{s \in [0, \ell]}$, where s denotes the arc length on the segment Γ_1 . Then the arc length along Γ_1 is equal to ℓ , and let L be the total arc length of the curve Γ . Given that the grid involved is uniform and Cartesian, we can expect that the average “step size” of the orthogonal projections of the discrete boundary γ onto the boundary curve Γ will be roughly the same on each segment, and it can be expressed as $L/|\gamma|$. We suggest as a rule-of-thumb that the minimum distance between roots of the largest Chebyshev polynomial $T_N(s)$ defined on Γ_1 (i.e., an interval of length ℓ) should not be larger

than the average resolution $L/|\gamma|$. In other words, we seek the largest N for which

$$\frac{L}{|\gamma|} \geq \min_{T_N(s)=0} (s_{j+1} - s_j) \quad (5.1)$$

is satisfied, where $s_{j+1} > s_j$ are consecutive roots of $T_N(s)$ on the interval $s \in [0, \ell]$. The roots of the Chebyshev basis function $T_N(x)$ on the canonical interval $x \in [-1, 1]$ are given by

$$x_j = \cos\left(\frac{2k-1}{2N}\pi\right), j = 1, \dots, N, \quad (5.2)$$

and it is clear from the symmetry of the cosine function that the minimum of (5.1) will always be achieved at the endpoints: x_1, x_2 or x_{N-1}, x_N . The interval $x \in [-1, 1]$ is mapped linearly to the arc length $s \in [0, \ell]$ so that $s = \frac{1}{2}(x\ell + \ell)$, see (3.32a). The minimum distance between roots on the interval $[0, \ell]$ in (5.1) then becomes

$$\min_{T_N(s)=0} (s_{j+1} - s_j) = s_2 - s_1 = \frac{1}{2}(x_2\ell + \ell) - \frac{1}{2}(x_1\ell + \ell) = \frac{\ell}{2}(x_2 - x_1). \quad (5.3)$$

Furthermore, by using the fourth-order Maclaurin approximation $\cos \theta \approx 1 - \frac{\theta^2}{2} + \mathcal{O}(\theta^4)$ we may approximate the difference $(x_2 - x_1)$ as follows:

$$x_2 - x_1 = \cos\left(\frac{\pi}{2N}\right) - \cos\left(\frac{3\pi}{2N}\right) \approx \frac{\pi^2}{N^2}. \quad (5.4)$$

Therefore, combining (5.4) with (5.3) and substituting into (5.1), the inequality (5.1) becomes

$$\frac{L}{|\gamma|} \geq \frac{\ell}{2} \frac{\pi^2}{N^2}, \quad (5.5)$$

which we rearrange to yield

$$N \leq \sqrt{\frac{|\gamma|\ell\pi^2}{2L}}. \quad (5.6)$$

The largest integer N which satisfies inequality (5.6) gives approximately the Chebyshev basis function for which the closest roots of the basis function on the interval $[0, \ell)$ are approximately equal to the average distance between the normal projections of the discrete boundary γ onto the continuous boundary curve Γ . As such, the bound (5.6) does not represent an exact bound for the largest number of basis functions that may be taken before accuracy is lost, which we illustrate by comparison to a previously explored case. It was noted in the discussion of Section 4.3.1 that for problems on a 64×64 grid with two equal boundary arcs of length π along the unit circle that the empirically determined saturation point for Chebyshev basis functions was around $N = 45$. By comparison, noting that on this grid the number of grid

nodes in the discrete boundary is $|\gamma| = 472$, the bound (5.6) becomes

$$N \leq \sqrt{\frac{|\gamma|\ell\pi^2}{2L}} = \sqrt{\frac{(472)(\pi)\pi^2}{2(2\pi)}} \approx 34. \quad (5.7)$$

5.2.2 Numerical investigation of the adequate truncation tolerance

In the following numerical investigation, we seek to gain insight into the relationship between the overall error of the scheme and the truncation tolerance σ of the series expansion of the boundary conditions. To do this, we will solve a problem on a single grid with what is known from previous experience to be more basis functions than needed to resolve the solution up to the full accuracy of the FD scheme for that problem. Then, we will solve the same problem again, removing basis functions one at a time. We repeat this until a detrimental effect on the error is observed. At that point, we compare the truncation error of the expansion of the boundary data with the overall error of the scheme. It is well known that the truncation error of the Chebyshev expansion can be estimated by the last coefficient in the expansion, and so the magnitude of the last Chebyshev coefficient will be used for comparison to the error of the scheme.

For this test we will solve the same mixed BC problem as in Section 3.3.4. This problem uses a split Chebyshev basis on the circular boundary Γ with segments $\Gamma_1 = \{(r, \theta) : r = 1, \theta \in [0, 2\pi/3]\}$ and $\Gamma_2 = \{(r, \theta) : r = 1, \theta \in [2\pi/3, 2\pi]\}$. The wavenumber for the problem is again $k = 10$. The only departure from the parameters specified in Section 3.3.4 will be that the number of basis functions used on each boundary segment will now be different from one another, as was introduced in Section 3.3.7. The computations are performed on a grid of size 512×512 , and we begin with a tolerance $\sigma = 3.8 \cdot 10^{-10}$, which exceeds the error for this grid found in Table 3.3 by several orders of magnitude. This results in bases of size $N_1 = 35$ on Γ_1 and $N_2 = 65$ on Γ_2 , which we reduce by 1 basis function on each segment for each trial. We denote by σ_1, σ_2 the empirical tolerances on the segments Γ_1, Γ_2 of the boundary curve. The results are summarized in Table 5.3.

We observe in Table 5.3 that accuracy is lost when the larger of the truncation tolerances becomes larger than the expected error, and this progression is clearly seen in the final three rows. This provides evidence that the tolerance σ on each boundary segment should be roughly equal to the error of the FD scheme in order to maintain accuracy of the overall algorithm.

5.3 A Numerical Example Using the *a priori* Error Estimate

In this example, we wish to use the average pollution coefficient $\overline{C_P} = 0.072$ obtained in Section 5.1.1 to solve a problem which shares the same geometry but is otherwise different from

Table 5.3: As the tolerances σ_1, σ_2 approach the error of the FD scheme, accuracy is lost.

(N_1, N_2)	Error	σ_1	σ_2
(35,65)	$4.12 \cdot 10^{-7}$	$3.80 \cdot 10^{-10}$	$3.80 \cdot 10^{-10}$
(34,64)	$4.12 \cdot 10^{-7}$	$1.14 \cdot 10^{-9}$	$7.69 \cdot 10^{-10}$
(33,63)	$4.11 \cdot 10^{-7}$	$3.33 \cdot 10^{-9}$	$1.52 \cdot 10^{-9}$
(32,62)	$4.11 \cdot 10^{-7}$	$9.62 \cdot 10^{-9}$	$2.98 \cdot 10^{-9}$
(31,61)	$4.11 \cdot 10^{-7}$	$2.74 \cdot 10^{-8}$	$5.83 \cdot 10^{-9}$
(30,60)	$4.12 \cdot 10^{-7}$	$7.70 \cdot 10^{-8}$	$1.13 \cdot 10^{-8}$
(29,59)	$4.10 \cdot 10^{-7}$	$2.13 \cdot 10^{-7}$	$2.18 \cdot 10^{-8}$
(28,58)	$5.30 \cdot 10^{-7}$	$5.79 \cdot 10^{-7}$	$4.19 \cdot 10^{-8}$
(27,57)	$1.28 \cdot 10^{-6}$	$1.55 \cdot 10^{-6}$	$7.99 \cdot 10^{-8}$

the problems solved thus far. Therefore, the domain Ω will continue to be a disk of radius 1 centered at the origin with the AP (2.51) which again has side length $s = 2.2$ as in all previous simulations of Chapters 3-4.

As a test solution, we take a sum of two plane waves in orthogonal directions and different wavelengths, $u = e^{ikx} + e^{i(2k)y}$. The wavenumber $k = 13.32$ is used since we have elsewhere used only integer values of k . Applying the Helmholtz operator L to this test solution results in an inhomogeneous right-hand side $f = -3k^2 e^{i(2k)y}$ of the BVP (3.1a). For the boundary condition (3.1b) on the disk, we take a Robin BC of the type (3.24) the following discontinuous Robin coefficients:

$$\begin{aligned}
 A(\theta) &= \begin{cases} \theta^2 + 1, & \theta \in [0, \pi) \\ \ln(\theta), & \theta \in [\pi, 2\pi) \end{cases}, \\
 B(\theta) &= \begin{cases} \cos \theta - \theta, & \theta \in [0, 3\pi/2) \\ \sin^2 \theta, & \theta \in [3\pi/2, 2\pi) \end{cases},
 \end{aligned} \tag{5.8}$$

and the right-hand side ϕ that is generated from plugging the test solution with these coefficients into (3.24). The series expansion of this Robin BC will require a 3-part Chebyshev basis on the segments $\Gamma_1 = \{r = 1, \theta \in [0, \pi)\}$, $\Gamma_2 = \{r = 1, \theta \in [\pi, 3\pi/2)\}$, $\Gamma_3 = \{r = 1, \theta \in [3\pi/2, 2\pi)\}$. In the cases where the number of basis functions determined by the tolerance σ exceed the maximum suggested on the grid by the inequality (5.6), the maximum is used, and this occurs on the grids of size 64×64 and 128×128 in Table 5.4. For the subsequent grids the tolerance $\sigma = \overline{C_P} h^4 k^5$ dictates a smaller number of basis functions than the maximum specified by (5.6), and the numbers of basis functions reported for these grids in Table 5.4 reflect the truncation as determined by this tolerance. The results are displayed in Table 5.4, and the real part of the numerically computed solution on the 256×256 grid is plotted in Figure 5.2.

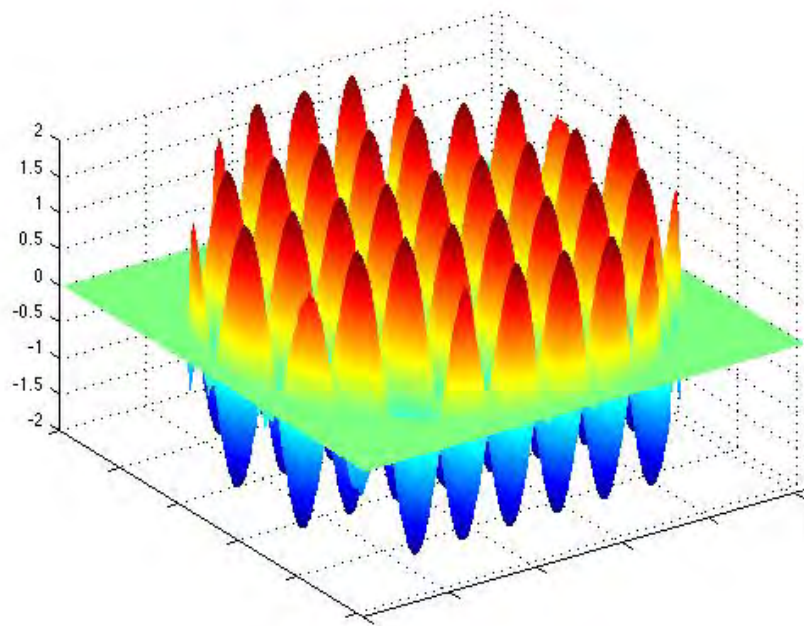


Figure 5.2: Real part of the numerically computed test solution $u = e^{ikx} + 2e^{i(2k)y}$ with wavenumber $k = 13.32$ on a 256×256 grid.

Table 5.4: The number of basis functions for the grids 64×64 and 128×128 is determined by the inequality (5.6), while for the finer grids it is determined by the truncation tolerance σ obtained from the average pollution coefficient from the previous computations of Table 5.2.

Grid	$\ u - u_{num}\ _\infty$	Convergence Rate	$\sigma = \overline{C_P} h^4 k^5$	(N_1, N_2, N_3)
64×64	4.97	-	$4.21 \cdot 10^{-2}$	(34,24,24)
128×128	$1.88 \cdot 10^{-2}$	8.05	$2.63 \cdot 10^{-3}$	(48,34,34)
256×256	$1.81 \cdot 10^{-4}$	6.70	$1.65 \cdot 10^{-4}$	(60,39,37)
512×512	$1.25 \cdot 10^{-5}$	3.85	$1.02 \cdot 10^{-5}$	(64,42,40)
1024×1024	$7.69 \cdot 10^{-7}$	4.02	$6.43 \cdot 10^{-7}$	(68,45,43)
2048×2048	$4.64 \cdot 10^{-8}$	4.05	$4.02 \cdot 10^{-8}$	(72,47,46)

It is noted that the overall error appears somewhat higher than our expectation for the coarser grids. This is due to the fact that we truncate the basis according to inequality (5.6) on each segment rather than using the tolerance σ since this could potentially result in a degradation of the linear independence of the basis functions at the discrete boundary γ due to the lack of sufficient resolution, as discussed in Section 5.2.1. As noted at the end of that discussion, it may be that we can safely take a few more basis functions and somewhat improve the overall accuracy, but, other than by experimentation, we do not know the actual saturation point of the basis. As the grid is refined by a factor of 2, the number of points at the grid boundary γ also roughly doubles and the lack of resolution is quickly alleviated, thus we are able to use the truncation tolerance σ to determine the number of basis functions and achieve a result that resembles the error estimate for the remaining grids, 256×256 and finer.

Chapter 6

Considerations for Future Work

This chapter summarizes a variety of ways in which the present work can be extended, and is organized as follows. In Section 6.1, direct extensions of the 2D algorithm presented in this dissertation are considered, while in Section 6.2 the more comprehensive extension to a 3D setting is discussed.

6.1 Extensions of the 2D Algorithm

Basis Functions One may note that we have employed two different choices of basis functions for the expansion of the boundary data in the preceding chapters: the Fourier and Chebyshev functions, see Sections 3.2.4 and 3.2.5, respectively. In general, any set of orthogonal basis functions may be used, and other choices may be desirable if, for example, they are expected to converge more rapidly for the problem at hand or are otherwise necessitated by the nature of the boundary conditions.

It should be noted then that if the boundary condition involves products of functions, such as the general variable coefficient Robin boundary condition (3.24), then our method relies on expressing the basis coefficients of a product as a convolution of the coefficients of individual factors, and formulae for such were used for both the Fourier and Chebyshev coefficients in this work — see formulae (3.25) and (A.2), respectively. The proposed methodology is valid for any system of basis functions on the boundary Γ for which a relation of this type can be obtained. However, if one desires to use a basis for which no corresponding formula can be conveniently obtained, an alternative to convolution-type formulae may be provided by collocation techniques, see, e.g., [42, 41].

Boundary Partitions One of the major contributions of this work was the piecewise parametrization of the boundary curve, as this is precisely what allowed for the treatment

of a wide range of boundary conditions, including mixed BCs (e.g., Section 3.3.4), Robin BCs with discontinuous coefficients (e.g., Section 3.3.5), and non-smooth boundary data (e.g., Section 4.3.2).

Partitioning the boundary into segments may alleviate other difficulties in addition to handling sophisticated boundary conditions. It may also be useful when, for example, the resolution at the boundary needs to be increased locally, which may be necessary on a more general boundary curve Γ on a segment with high curvature. Another potential use is when the boundary itself is defined as a composition of independently parameterized segments rather than as one curve with a global parametrization, such as a square or other shapes with corners.

A partitioned boundary may also be used in cases where there is no reason inherent to the geometry or BC to do so, as was demonstrated in, e.g., Section 3.3.8. This may be useful if one desires to solve several related problems which may or may not have complications in the boundary conditions at known locations along the boundary curve, so that one and the same partition of the boundary curve (and therefore the same basis) may be used either way, the efficiency of which we have remarked on in Sections 3.3.1 and 4.2.2.

Singularities In Chapter 4, we implement the method of singularity subtraction with a local conformal mapping to treat singularities resulting from discontinuous boundary data while maintaining the high-order accuracy of the scheme. The case of singularities arising from geometric corners is remarked upon in Section 4.1.1, but was only used as a foundation for curved boundaries via the conformal mapping in Section 4.1.2. A logical next step would be to consider near-boundary singularities that are due not only to the discontinuities in the data, but also to geometric irregularities, such as corners or cusps. All of the necessary analysis for this formulation has already been performed in Section 4.1. To make the overall approach more general, one will also need to account for a larger class of boundary conditions beyond the Dirichlet and Neumann cases treated in this work. An even more comprehensive extension would involve the analysis of singularities at the interface between two materials when solving transmission/scattering problems, see [17], wherein singularities may arise from discontinuities in the interface conditions/data and/or the geometric irregularities of the interface itself.

Higher Wavenumbers We also note that neither in this dissertation nor in previous papers devoted to solving the Helmholtz equation by the method of difference potentials, see [25, 18, 17], has the performance of this method been extensively studied in the case of large wavenumbers. Section 5.1.1 presents the highest wavenumbers yet used with the method of difference potentials, with the largest being $k = 26$. Although results for the wavenumbers $k = 3, \dots, 26$ are presented there in Table 5.2, the design of the numerical study and subsequent analysis only begins to shed light on the subject. As remarked upon in Section 5.2, larger wavenumbers

will require higher dimensions N of the basis chosen on the boundary Γ , see formula (3.17), increasing the overall computational load. We expect, however, that N will increase slowly since the solutions we are computing are sufficiently smooth by design, even for problems with boundary singularities (since the singularities are removed prior to numerical approximation). The choice of the tolerance σ for the truncation of the series expansion of the boundary data will also be vital to maintaining efficiency, see Section 5.2. The topic of higher wavenumbers may require additional consideration in the future, but in the meantime we mention only that, similarly to any other finite difference approach, our methodology is prone to the pollution effect, which is precisely the reason why we use high order accurate schemes, see Chapter 2.

The obvious way to combat the pollution effect as the wavenumber becomes very large is by using even higher order schemes; however, one may note that a higher order compact scheme will require a higher order Taylor extension (see Section 3.2.1). Sixth-order accurate schemes for the Helmholtz equation have been developed for constant-coefficient [31] as well as variable wavenumber [32] equations. Increasing the order of accuracy even further naturally leads to a question of whether a spectral approximation can be used. On one hand, the AP is always formulated on a simple domain, such as a rectangle. This is done primarily for the reason of making its numerical solution easy and efficient, and of course, such an AP can be as easily solved by a spectral method as by high order finite differences. On the other hand, in the core of the method of difference potentials is the reduction of the governing equation from the domain to the boundary in the form of the BEP (3.4). This requires that the equation be approximated on a local stencil, so that the grid boundary γ is located near Γ . For spectral methods this is not intuitively possible because formally one can interpret a spectral discretization as having a stencil that occupies the entire domain. Altogether, we leave the question of further improving the approximation accuracy beyond 4th or 6th order in the method of difference potentials for future investigation.

6.2 Extension of the algorithm to 3D

Previously, the method of difference potentials has been used in 3D to construct the artificial boundary conditions for fluid flow computations [66] and for quasi-state plasma simulations [15], but no work of the nature that we propose. The methods presented in this dissertation can be extended to problems in 3D; however, it will require building and testing all the basic components of an algorithm based on the method of difference potentials. While all the key ideas of the method from 2D remain the same, its implementation for solving 3D boundary value problems will imply a number of substantial changes throughout the entire procedure, and many of these issues which require special attention are due to the geometry.

Boundary surfaces In 3D, it is more challenging to determine whether a given point (i.e., a grid node) lies inside or outside of a given closed surface than for the analogous 2D problem. An efficient solution to this problem determines the efficiency of partitioning of the set \mathbb{M}_0 into \mathbb{M}^+ and \mathbb{M}^- , see Section 3.1.2 and Figure 3.1. This problem has been extensively studied in the areas of computer graphics and computational geometry [67, 68, 69, 70] and subsequent work on adaptive Cartesian mesh generation for local grid refinement was developed as part of Cart3D [71] even for moving boundaries in time-dependent problems [72].

Finding a convenient representation for the boundary surface Γ may also prove more difficult than in 2D (unless it is a simple analytical shape, e.g., a sphere) and may require using multiple patches to parameterize the wide class of shapes in which we are interested (this approach has been used in [73, 74]). Furthermore, choosing a good basis (3.16) on the two-dimensional surface Γ that would enable an efficient (i.e., low-dimensional) representation (3.17) may be more difficult than constructing a basis on the one-dimensional curve Γ . In 2D, Fourier or Chebyshev bases were sufficient to represent one-dimensional functions on the boundary curve, but choosing a suitable basis for two-dimensional functions on the boundary surface which admit an efficient representation may be more difficult. Construction of the extension operator $\mathbf{E}\mathbf{x}$, see Section 3.2.1, is likely to require special surface-oriented coordinates [75], whereas in 2D this can be done in the general case by representing a curve by normal and tangential coordinates.

Solvers Finite difference schemes for the 3D Helmholtz equation will also be needed. A fourth order method for the variable-coefficient Helmholtz equation in 3D may be developed by the same equation-based method of Chapter 2 for 2D. In fact, sixth order schemes for the Helmholtz equation with variable wavenumber in both 2D and 3D has been introduced and tested in [32].

However, unlike in 2D, preconditioned iterative solvers provide the only realistic avenue for solving the AP in the case of variable coefficients, as opposed to the direct solvers that we have used in 2D. An iterative solver that can be easily parallelized to the Helmholtz equation discretized by compact high-order schemes has been successfully implemented in [32]. Additionally, a class of efficient complex-shifted Helmholtz preconditioners that can be inverted by multigrid is described in [76, 77]. For constant coefficient interior problems, the solution by the separation of variables method, as was suggested for 2D in Section 3.1.1, will remain the most efficient approach in 3D; however, for exterior problems one may use convolution with the discrete fundamental solution, which automatically takes into account the proper behavior of the solution in the far field, accelerated by the fast multipole method, see [15, Appendix C].

REFERENCES

- [1] A. Bayliss, C.I. Goldstein, and E. Turkel. On accuracy conditions for the numerical computation of waves. *Journal of Computational Physics*, 59(3):396–404, 1985.
- [2] I.M. Babuška and S.A. Sauter. Is the pollution effect of the fem avoidable for the helmholtz equation considering high wave numbers? *Siam Review*, 42:451–484, 2000.
- [3] I.M. Babuška, P. Bouillard, and A. Deraemaeker. Dispersion and pollution of the FEM solution for the Helmholtz equation in one, two and three dimensions. *Int. J. Numer. Meth. Engnr.*, 48:471–499, 1999.
- [4] A. C. Cangellaris and D.B. Wright. Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena. *IEEE Trans. Antennas and Propagation*, 39:1518–1525, 1991.
- [5] R. Holland. Pitfalls of staircase meshing. *IEEE Trans. on Electromagnetic Compatibility*, 35(4):434–439, 1993.
- [6] A.P. Calderon. Boundary-value problems for elliptic equations. In *Proceedings of the Soviet-American Conference on Partial Differential Equations*, pages 303–304, Novosibirsk, Moscow, Russia, 1963. Fizmatgiz.
- [7] R.T. Seeley. Singular integrals and boundary value problems. *Amer. J. Math.*, 88:781–809, 1966.
- [8] V.S. Ryaben’kii. *Method of Difference Potentials and Its Applications*, volume 30 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2002.
- [9] V.S. Ryaben’kii. Boundary equations with projections. *Russian Math. Surveys*, 40:147–183, 1985.
- [10] V.S. Ryaben’kii. Difference potentials method and its applications. *Math. Nachr.*, 177:251–264, 1996.

- [11] V.S. Ryaben'kii. On the method of difference potentials. *J. Sci. Comput.*, 28:467–478, 2006.
- [12] J. Lončarić, V.S. Ryaben'kii, and S.V. Tsynkov. Active shielding and control of noise. *SIAM J. Appl. Math.*, 62:565–596, 2001.
- [13] S.V. Tsynkov. On the definition of surface potentials for finite-difference operators. *J. Sci. Comp*, 18:155–189, 2003.
- [14] S.V. Utyuzhnikov. Generalized Calderón-Ryaben'kii's potentials. *IMA J. Appl. Math.*, 74:128–148, 2009.
- [15] E. Kansa, U. Shumlak, and S.V. Tsynkov. Discrete Calderon's projections on parallelepipeds and their application to computing exterior magnetic fields for FRC plasmas. *J. Comput. Phys.*, 234:172–198, 2013.
- [16] V.S. Ryaben'kii. Difference potentials analogous to Cauchy integrals. *Russian Math. Surveys*, 67:541–567, 2012.
- [17] M. Medvinsky, S.V. Tsynkov, and E. Turkel. High order numerical simulation of the transmission and scattering of waves using the method of difference potentials. *J. Comput. Phys.*, 243:305–322, 2013.
- [18] M. Medvinsky, S.V. Tsynkov, and E. Turkel. The method of difference potentials for the Helmholtz equation using compact high order schemes. *J. Sci. Comput.*, 53:150–193, 2012.
- [19] L. Fox and R. Sankar. Boundary singularities in linear elliptic differential equations. *J. Inst. Math. Appl.*, 5:340–350, 1969.
- [20] D.S. Kamenetskii. A numerical method for solving a singular boundary value problem for the chaplygin equation in the hodograph plane. *Preprint*, 1992.
- [21] D.S. Kamenetskii and V.S. Ryaben'kii. Solution of boundary value problems for the laplace equation in a domain with a cut by the method of difference potentials. *Preprint*, 1990.

- [22] D.S. Kamenetskii and E.G. Shifrin. Design of nozzle blades of a gas turbine by the method of hodograph. *Preprint 61*, 1992.
- [23] S. Britt, S.V. Tsynkov, and E. Turkel. A compact fourth order scheme for the Helmholtz equation in polar coordinates. *J. of Sci. Comp.*, 45:1–3, 2010.
- [24] S. Britt, S.V. Tsynkov, and E. Turkel. Numerical simulation of time-harmonic waves in inhomogeneous media using compact high order schemes. *Commun. Comput. Phys.*, 9:520–541, 2011.
- [25] S. Britt, S.V. Tsynkov, and E. Turkel. A high order numerical method for the Helmholtz equation with non-standard boundary conditions. *SIAM J. Sci. Comput.*, 35:A2255–A2292, 2013.
- [26] S. Britt, S. Petropavlovsky, S.V. Tsynkov, and E. Turkel. Computation of singular solutions to the helmholtz equation with high order accuracy. *Accepted by App. Num. Math.*, 2014.
- [27] I. Harari and E. Turkel. Difference methods for time-harmonic wave propagation. *J. Comput. Phys.*, 119:252–270, 1995.
- [28] I. Singer and E. Turkel. High-order finite difference methods for the Helmholtz equation. *Comput. Methods Appl. Mech. Engrg.*, 163:343–356, 1998.
- [29] P.D. Lax and R.D. Richtmyer. Survey of the stability of linear finite difference equations. *Commun. on Pure and App. Math.*, IX:267–293, 1956.
- [30] I. Singer and E. Turkel. High-order finite difference methods for the Helmholtz equation. *Comput. Methods Appl. Mech. Engrg.*, 163:343–358, 1998.
- [31] I. Singer and E. Turkel. Sixth-order accurate finite difference schemes for the Helmholtz equation. *J. Comput. Acoust.*, 14:339–351, 2006.

- [32] D. Gordon, R. Gordon, E. Turkel, and S.V. Tsynkov. Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number. *J. Comput. Phys.*, 232:272–287, 2013.
- [33] G. Baruch, G. Fibich, and S.V. Tsynkov. A high-order numerical method for the nonlinear Helmholtz equation in multidimensional layered media. *J. Comput. Phys.*, 228:3789–3815, 2009.
- [34] Y. Erlangga and E. Turkel. Iterative schemes for high order compact discretizations to the exterior Helmholtz equation. *ESAIM: M2AN*, 46:647–660, 2012.
- [35] V.I. Smirnov. *Course of Higher Mathematics. Vol. IV. Part 2.* Nauka, Moscow, Sixth revised edition, 1981.
- [36] V.I. Smirnov. *Course of Higher Mathematics. Vol. IV. Part 1.* Nauka, Moscow, Sixth revised edition, 1974.
- [37] E. Turkel. Comments on iterative schemes for high order compact discretizations to the exterior Helmholtz equation. *ESAIM: M2AN*, 49:221–223, 2015.
- [38] A.A. Reznik. Approximation of surface potentials of elliptic operators by difference potentials. *Dokl. Akad. Nauk SSSR*, 263:1318–1321, 1982.
- [39] A.A. Reznik. *Approximation of the surface potentials of elliptic operators by difference potentials and solution of boundary-value problems.* PhD thesis, Moscow Institute for Physics and Technology, Moscow, USSR, 1983.
- [40] V.S. Ryaben’kii and S.V. Tsynkov. *A Theoretical Introduction to Numerical Analysis.* Chapman & Hall/CRC, Boca Raton, FL, 2007.
- [41] J.S. Hesthaven, D. Gottlieb, and S. Gottlieb. *Spectral methods for time-dependent problems*, volume 21 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2007.

- [42] J.P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover Publications Inc., Mineola, NY, 2nd edition, 2001.
- [43] H. Tal-Ezer. Non-periodic trigonometric polynomial approximation. *J. Sci. Comput.*, 60:345–362, 2014.
- [44] A. Bayliss and E. Turkel. Mappings and accuracy for Chebyshev pseudo-spectral approximations. *J. Comput. Phys.*, 101:349–359, 1992.
- [45] G.C. Gaunaurd and H.C. Strifors. Scattering of electromagnetic pulses by simple-shaped targets with radar cross section modified by a dielectric coating. *IEEE Transactions on Antennas and Propagation*, 46:1252–1262, 1998.
- [46] M.L. Levin, S.M. Rytov, and V.D. Shafranov. M. A. Leontovich’s researches in electrodynamics. *Sov. Phys. Usp.*, 26:353–355, 1983.
- [47] H.H. Syed and J.L. Volakis. Electromagnetic scattering by coated convex surfaces and wedges simulated by approximate boundary conditions. In *NASA Contractor Report 190540*, 1992.
- [48] T.B.A. Senior and J.L. Volakis. *Approximate Boundary Conditions in Electromagnetics*, volume 41 of *IEEE Electromagnetic Waves Series*. IEEE, London, UK, 1995.
- [49] C.S. Peskin. The immersed boundary method. *Acta Numer.*, 11:479–517, 2002.
- [50] K. Ito and Z. Li. The immersed interface method. In *SIAM Frontiers in App. Math.*, volume 33, Philadelphia, PA, 2006.
- [51] T. Aslam, R.P. Fedkiw, and S. Xu. The ghost fluid method for deflagration and detonation discontinuities. *J. Comput. Phys.*, 154:393–427, 1999.
- [52] R.P. Fedkiw, M. Kang, and X.D. Liu. A boundary condition capturing method for Poisson’s equation on irregular domains. *J. Comput. Phys.*, 160:151–178, 2000.

- [53] P. Colella and H. Johansen. A cartesian grid embedded boundary method for poisson's equation on irregular domains. *J. Comput. Phys.*, 147(1):60–85, 1998.
- [54] P. Colella, R.K. Crockett, and D.T. Graves. A Cartesian grid embedded boundary method for solving the Poisson and heat equations with discontinuous coefficients in three dimensions. *J. Comput. Phys.*, 230:2451–2469, 2011.
- [55] F. Oberhettinger. Diffraction of waves by a wedge. *Comm. Pure Appl. Math.*, 7:551–563, 1954.
- [56] W. Wasow. Asymptotic development of the solution of Dirichlet's problem at analytic corners. *Duke Math. J.*, 24:47–56, 1957.
- [57] R.S. Lehman. Developments at an analytic corner of solutions of elliptic partial differential equations. *J. Math. Mech.*, 8:727–760, 1959.
- [58] N.M. Wigley. Asymptotic expansions at a corner of solutions of mixed boundary value problems. *J. Math. Mech.*, 13:549–576, 1964.
- [59] R.J. Riddell Jr. Boundary-distribution solution of the Helmholtz equation for a region with corners. *J. Comput. Phys.*, 31:21–41, 1979.
- [60] R.J. Riddell Jr. Numerical solution of the Helmholtz equation for two-dimensional polygonal regions. *J. Comput. Phys.*, 31:42–59, 1979.
- [61] M. Costabel and M. Dauge. Construction of corner singularities for Agmon-Douglis-Nirenberg elliptic systems. *Math. Nachr.*, 162:209–237, 1993.
- [62] J. Berger, V. Mantič, and F. París. Singularities in 2D anisotropic potential problems in multi-material corners: real variable approach. *Internat. J. Solids Structures*, 40:5197–5218, 2003.
- [63] D. Lesnic, V. Mantič, and L. Marin. Treatment of singularities in Helmholtz-type equations using the boundary element method. *J. Sound Vibration*, 278:39–62, 2004.

- [64] L. Marin. Treatment of singularities in the method of fundamental solutions for two-dimensional Helmholtz-type equations. *Appl. Math. Model.*, 34:1615–1633, 2010.
- [65] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.
- [66] S.V. Tsynkov. External boundary conditions for three-dimensional problems of computational aerodynamics. *SIAM J. Sci. Comp.*, 21:166–206, 1999.
- [67] S. Feiner, J. Foley, J. Hughes, and A. van Dam. *Computer Graphics: Principles and Practice*. Addison-Wesley, Reading, MA, 1995.
- [68] J. ORourke. *Computational Geometry*. Cambridge Univ. Press, NY, 1993.
- [69] F.P. Preparata and M.I. Shamos. *Computational Geometry: An Introduction*. Springer-Verlag, 1985.
- [70] D. Voorhies. *Graphics Gems II: Triangle-Cube Intersections*. Academic Press, Inc., 1992.
- [71] M.J. Aftosmis, M.J. Berger, and J.E. Melton. Adaptive cartesian mesh generation, 1998.
- [72] M.J. Aftosmis, M.J. Berger, and S.M. Murman. Implicit approaches for moving boundaries in a 3-d cartesian method. In *41st AIAA Aerospace Sciences Meeting*, 2003.
- [73] O.P. Bruno and L.A. Kunyansky. Surface scattering in three dimensions: an accelerated high-order solver. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 457:2921–2934, 2001.
- [74] O.P. Bruno and L.A. Kunyansky. A fast, high-order algorithm for the solution of surface scattering problems: basic implementation, tests, and applications. *J. Comput. Phys.*, 169:80–110, 2001.
- [75] E.H. Hirschel and W. Kordulla. *Shear flow in surface-oriented coordinate*, volume 4. Vieweg, 1981.
- [76] Y.A. Erlangga, C.W. Oosterlee, and C. Vuik. On a class of preconditioners for solving the Helmholtz equation. *Appl. Numer. Math.*, 50:409–425, 2004.

- [77] Y. Erlangga and E. Turkel. Iterative schemes for high order compact discretizations to the exterior Helmholtz equation. *ESAIM: M2AN*, 46:647–660, 2012.

APPENDICES

Appendix A

Chebyshev coefficients for the product of two functions

The expansion of a smooth bounded function $f(x)$, $x \in [-1, 1]$, in the Chebyshev basis $\{T_n(x)\}_{n=0}^{\infty}$ is given by

$$f(x) = \sum_{n=0}^{\infty} \hat{f}_n T_n(x), \quad \text{where} \quad \hat{f}_n = \begin{cases} \frac{1}{2} \int_{-1}^1 \omega(x) f(x) T_n(x) dx, & n = 0, \\ \int_{-1}^1 \omega(x) f(x) T_n(x) dx, & n > 0. \end{cases} \quad (\text{A.1})$$

Let $f(x)$ and $g(x)$ be smooth on $x \in [-1, 1]$. Then, according to the definition of the Chebyshev coefficients, see (A.1), we have:

$$\begin{aligned} (\widehat{fg})_0 &= \frac{1}{2} \int_{-1}^1 \omega(x) f(x) g(x) T_0(x) dx = \frac{1}{2} \int_{-1}^1 \omega(x) f(x) \left[\sum_{m=0}^{\infty} \hat{g}_m T_m \right] dx \\ &= \frac{1}{2} \sum_{m=0}^{\infty} \hat{g}_m \int_{-1}^1 \omega(x) f(x) T_m(x) dx = \frac{1}{2} \left(\hat{g}_0 (2\hat{f}_0) + \sum_{m=1}^{\infty} \hat{g}_m \hat{f}_m \right) = \hat{g}_0 \hat{f}_0 + \frac{1}{2} \sum_{m=1}^{\infty} \hat{g}_m \hat{f}_m. \end{aligned} \quad (\text{A.2a})$$

For $n > 0$, we take into account that

$$T_n(x) T_m(x) = \begin{cases} \frac{1}{2} (T_{m+n}(x) + T_{m-n}(x)), & m \geq n, \\ \frac{1}{2} (T_{m+n}(x) + T_{n-m}(x)), & m < n, \end{cases}$$

and using (A.1) obtain:

$$\begin{aligned}
(\widehat{fg})_n &= \int_{-1}^1 \omega(x) f(x) g(x) T_n(x) dx = \int_{-1}^1 \omega(x) f(x) \left[\sum_{m=0}^{\infty} \hat{g}_m T_m \right] T_n(x) dx \\
&= \sum_{m=0}^{\infty} \hat{g}_m \int_{-1}^1 \omega(x) f(x) T_m(x) T_n(x) dx \\
&= \sum_{m=0}^{n-1} \hat{g}_m \cdot \frac{1}{2} \left[\left(\int_{-1}^1 \omega(x) f(x) T_{m+n}(x) dx \right) + \left(\int_{-1}^1 \omega(x) f(x) T_{n-m}(x) dx \right) \right] \quad (\text{A.2b}) \\
&\quad + \sum_{m=n}^{\infty} \hat{g}_m \cdot \frac{1}{2} \left[\left(\int_{-1}^1 \omega(x) f(x) T_{m+n}(x) dx \right) + \left(\int_{-1}^1 \omega(x) f(x) T_{m-n}(x) dx \right) \right] \\
&= \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m (\hat{f}_{m+n} + \hat{f}_{n-m}) + \hat{g}_n (\hat{f}_{2n} + 2\hat{f}_0) + \sum_{m=n+1}^{\infty} \hat{g}_m (\hat{f}_{m+n} + \hat{f}_{m-n}) \right).
\end{aligned}$$

We expect formula (A.2b) to be symmetric with respect to f and g . While this is not immediate in the formula itself, it can easily be shown. To see the symmetry of (A.2b) with respect to f and g , first re-arrange the summations by the \hat{f} subscripts (i.e., one summation for $m+n$ terms, one for $n-m$, and one for $m-n$). Then, after appropriate substitutions in the summation indices (respectively, $j = m+n$, $j = n-m$, and $j = m-n$), the form of (A.2b) with interchanged g and f terms can be obtained by regrouping the summation terms.

Appendix B

Singularity Subtraction Test Problems 2 through 6

Test 2 The Dirichlet boundary data for this test have discontinuous first derivative at $(\pm R, 0)$:

$$u|_{r=R} = \begin{cases} \pi/2 - \varphi, & 0 < \varphi < \pi, \\ \varphi - 3\pi/2, & \pi < \varphi < 2\pi, \end{cases}$$

or, after the mapping (4.19),

$$u(\xi, 0) = \begin{cases} \frac{\pi}{2} + \arctan \frac{2\xi}{\xi^2 - 1}, & \xi > 0, \\ \frac{\pi}{2} - \arctan \frac{2\xi}{\xi^2 - 1}, & \xi < 0. \end{cases}$$

Equivalently, in polar coordinates (ρ, θ) on the (ξ, η) plane on the edges of the wedge with angle π we have:

$$\begin{aligned} u|_{\theta=0} &= \frac{\pi}{2} + \arctan \frac{2\rho}{\rho^2 - 1}, \\ u|_{\theta=\pi} &= \frac{\pi}{2} + \arctan \frac{2\rho}{\rho^2 - 1}. \end{aligned} \tag{B.1}$$

The Taylor expansion of the data (B.1) at $\rho = 0$ (which corresponds to the singular point $(R, 0)$) reads:

$$u|_{\theta=0,\pi} = \frac{\pi}{2} - 2\rho + \frac{2}{3}\rho^3 - \frac{2}{5}\rho^5 + \dots \tag{B.2}$$

Note that the formally identical expressions (B.1), (B.2) for the boundary conditions in the coordinates (ρ, θ) undergo a jump in the derivative since the differentiation w.r.t. ρ is done in the opposite directions on the rays $\theta = 0$ and $\theta = \pi$. Moreover, after the conformal mapping the singularity of the boundary data appears in all odd derivatives.

The coefficients of the asymptotic expansion at the point $(R, 0)$ are given by

$$\begin{aligned} A_0(\theta) = A_2(\theta) = 0, \quad A_1(\theta) &= \frac{4 \sin \theta}{\pi}, \\ A_3(\theta) &= \frac{6 k^2 R^2 - 4}{3\pi} \sin 3\theta - \frac{2 k^2 R^2}{\pi} \sin \theta, \quad A_4(\theta) = \frac{4 k^2 R^2 (\cos 2\theta - 1)^2}{3\pi}, \end{aligned} \quad (\text{B.3})$$

and

$$\begin{aligned} B_0(\theta) &= \frac{\pi}{2}, \quad B_1(\theta) = C_1^{(B)} \sin \theta + \frac{2(2\theta - \pi) \cos \theta}{\pi}, \\ B_2(\theta) &= C_2^{(B)} \sin 2\theta + \frac{\pi k^2 R^2}{2} (\cos 2\theta - 1), \\ B_3(\theta) &= -\frac{1}{2} k^2 R^2 \sin \theta C_1^{(B)} + \sin 3\theta C_3^{(B)} + \\ &\quad \frac{1}{9\pi} \left[12 \{ (6k^2 R^2 - 4) \cos^2 \theta - 6k^2 R^2 + 3 \} (\theta - \pi/2) \cos \theta + \right. \\ &\quad \left. \sin \theta \{ 9k^2 R^2 (\pi^2 + 4/3) - 2 \} + \sin \theta \cos^2 \theta (8 - 12k^2 R^2) \right], \\ B_4(\theta) &= \frac{1}{3} k^2 R^2 (\cos 2\theta - 1)^2 C_1^{(B)} - \frac{1}{3} k^2 R^2 \sin 2\theta C_2^{(B)} + \sin 4\theta C_4^{(B)} - \\ &\quad \frac{k^2 R^2}{6\pi} \left[\cos^2 2\theta + \{ 8\theta \sin 2\theta - 16/3 + (k^2 R^2 - 6) \pi^2 \} \cos 2\theta \right. \\ &\quad \left. + \{ \pi^2 (3 - k^2 R^2/4) - 7/6 \} \cos 4\theta + (8\pi - 16\theta) \sin 2\theta + 11/2 + 3\pi^2 (1 - k^2 R^2/4) \right]. \end{aligned} \quad (\text{B.4})$$

The asymptotic expansion at the opposite singular point $(-R, 0)$ is obtained from that at the point $(R, 0)$ by an odd reflection about the y axis using the symmetry of the boundary condition.

Test 3 The Dirichlet boundary data for this test have discontinuous second derivative (as well as higher order even derivatives) at the points $(\pm R, 0)$:

$$u|_{r=R} = \begin{cases} \cos \varphi, & 0 < \varphi < \pi, \\ \cos 3\varphi, & \pi < \varphi < 2\pi. \end{cases}$$

After the conformal mapping we have:

$$u(\xi, 0) = \begin{cases} \cos \left(\arctan \frac{2\xi}{\xi^2 - 1} \right), & \xi > 0, \\ \cos \left(3 \arctan \frac{2\xi}{\xi^2 - 1} \right), & \xi < 0. \end{cases} \quad (\text{B.5})$$

Finally, in terms of (ρ, θ) the boundary condition (B.5) translates into:

$$\begin{aligned} u|_{\theta=0} &= \cos\left(\arctan\frac{2\rho}{\rho^2-1}\right) \underset{\rho \rightarrow 0}{\approx} 1 - 2\rho^2 + 2\rho^4 - 2\rho^6 + \dots, \\ u|_{\theta=\pi} &= \cos\left(3\arctan\frac{2\rho}{\rho^2-1}\right) \underset{\rho \rightarrow 0}{\approx} 1 - 18\rho^2 + 66\rho^4 - 146\rho^6 + \dots \end{aligned} \quad (\text{B.6})$$

The coefficients of the asymptotic expansion at the point $(R, 0)$ are:

$$\begin{aligned} A_0(\theta) &= A_1(\theta) = A_3(\theta) \equiv 0, \quad A_2 = -\frac{16}{\pi} \sin 2\theta, \\ A_4(\theta) &= \frac{16}{3\pi} [(12 - k^2 R^2) \sin 4\theta + k^2 R^2 \sin 2\theta], \end{aligned} \quad (\text{B.7})$$

$$\begin{aligned} B_0(\theta) &= 1, \quad B_1(\theta) = 0, \\ B_2(\theta) &= C_2^{(B)} \sin 2\theta + \left(k^2 R^2 - 2 - \frac{16}{\pi} \theta\right) \cos 2\theta + \frac{8}{\pi} \sin 2\theta - k^2 R^2, \\ B_3(\theta) &= \sin 3\theta C_3^{(B)} + 2k^2 R^2 \sin \theta, \\ B_4(\theta) &= -\frac{1}{3} k^2 R^2 \sin 2\theta C_2^{(B)} + \sin 4\theta C_4^{(B)} + \left(\frac{k^4 R^4}{12} - \frac{5k^2 R^2}{3} + 2\right) \cos 4\theta \\ &\quad + \frac{1}{36\pi} \left[(4608 - 384k^2 R^2) \theta \cos^2 2\theta + \{ (48k^2 R^2 - 576) \sin 2\theta \right. \\ &\quad \left. - 12k^2 R^2 (\pi k^2 R^2 - 8\pi - 16\theta) \} \cos 2\theta - 160k^2 R^2 \sin 2\theta + 9\pi k^4 R^4 \right. \\ &\quad \left. + 192k^2 R^2 (\theta - 3\pi/16) - 2304\theta \right]. \end{aligned} \quad (\text{B.8})$$

Similarly to Test 2, the boundary data are anti-symmetric w.r.t. the y axis, and thus the asymptotic expansion at $(-R, 0)$ is given by an odd reflection of its counterpart at the point $(R, 0)$.

Test 4 This is a Neumann test problem, for which the boundary conditions are set for the normal derivative of the solution. We start with the case of a discontinuity in the normal derivative itself:

$$\left. \frac{\partial u}{\partial r} \right|_{r=R} = \begin{cases} 1, & 0 < \varphi < \pi, \\ 0, & \pi < \varphi < 2\pi. \end{cases}$$

Obviously, the normal derivative undergoes a unit jump in the circumferential direction at the points $(\pm R, 0)$ (cf. formula (4.28)). It can be easily verified that under the mapping (4.19) the normal derivative at the boundary in the (ξ, η) coordinates takes the form:

$$\left. \frac{\partial u}{\partial r} \right|_{r=R} = -\frac{1}{2R} (1 + \xi^2) \left. \frac{\partial u}{\partial \eta} \right|_{\eta=0},$$

where

$$\left. \frac{\partial u}{\partial \eta} \right|_{\eta=0} = \begin{cases} -\frac{2R}{1+\xi^2}, & \xi > 0, \\ 0, & \xi < 0, \end{cases}$$

and the corresponding (ρ, θ) expressions are

$$\begin{aligned} \left. \frac{1}{\rho} \frac{\partial u}{\partial \theta} \right|_{\theta=0} &= -\frac{2R}{1+\rho^2}, \\ \left. \frac{1}{\rho} \frac{\partial u}{\partial \theta} \right|_{\theta=\pi} &= 0. \end{aligned} \tag{B.9}$$

Note that in general the following relation holds between the derivatives:

$$\left. \frac{\partial u}{\partial \eta} \right|_{\eta=0} = -\left. \frac{1}{\rho} \frac{\partial u}{\partial \theta} \right|_{\theta=\pi}, \quad \xi < 0. \tag{B.10}$$

Hence, one must remember the additional minus sign on the negative semi-axis $\xi < 0$. For this particular test problem though, this issue is obscured by the zero condition on $\xi < 0$.

Finally, the power series expansion of the conditions (B.9) on the rays $\theta = 0, \theta = \pi$ is

$$\begin{aligned} \left. \frac{\partial u}{\partial \theta} \right|_{\theta=0} &= -\frac{2R\rho}{1+\rho^2} \underset{\rho \rightarrow 0}{\approx} -2R(\rho - \rho^3 + \rho^5 - \dots), \\ \left. \frac{\partial u}{\partial \theta} \right|_{\theta=\pi} &= 0. \end{aligned}$$

The coefficients of expansion (4.16) at the point $(R, 0)$ are given by

$$\begin{aligned} A_0(\theta) &= -\frac{2R \cos \theta}{R}, \quad A_1(\theta) = 0, \\ A_2(\theta) &= \frac{R(2 - k^2 R^2)}{3\pi} \cos 3\theta + \frac{k^2 R^3}{\pi} \cos \theta, \quad A_3(\theta) = \frac{2k^2 R^3}{3\pi} (\sin 4\theta - 2 \sin 2\theta), \end{aligned} \tag{B.11}$$

$$\begin{aligned} B_0(\theta) &= C_0^{(B)} \cos \theta + \frac{2R}{\pi} [(\theta - \pi) \sin \theta + \cos \theta], \quad B_1(\theta) = C_1^{(B)} \cos 2\theta, \\ B_2(\theta) &= -\frac{k^2 R^2}{2} \cos \theta C_0^{(B)} + C_3^{(B)} \cos 3\theta + \frac{4R}{3\pi} \left[(k^2 R^2 / 6 - 1/3) \cos^3 \theta \right. \\ &\quad \left. - (5k^2 R^2 / 4 - 1/4) \cos \theta - \frac{1}{2} \sin \theta (\pi - \theta) \{ 2 \cos^2 \theta (k^2 R^2 - 2) + 1 - 2k^2 R^2 \} \right], \tag{B.12} \\ B_3(\theta) &= \frac{k^2 R^2}{3} (2 \sin 2\theta - \sin 4\theta) C_0^{(B)} + C_3^{(B)} \cos 4\theta + \frac{4k^2 R^3}{3\pi} \left[\theta \{ \cos^2 2\theta + 1/4 \} \right. \\ &\quad \left. + \left\{ \pi - \theta - \frac{1}{8} \sin 2\theta \right\} \cos 2\theta + \frac{4}{3} \sin 2\theta - \frac{2}{3} \sin 4\theta - \frac{3}{4} \pi \right]. \end{aligned}$$

The asymptotic expansion at $(-R, 0)$ is obtained by an even reflection about the y axis.

Test 5 In this test, the Neumann boundary data undergo a jump discontinuity in the first derivative,

$$\left. \frac{\partial u}{\partial r} \right|_{r=R} = \begin{cases} \pi/2 - \varphi, & 0 < \varphi < \pi, \\ \varphi - 3\pi/2, & \pi < \varphi < 2\pi. \end{cases}$$

In terms of (ρ, θ) , we have (in accordance with the comment right after equation (B.10)):

$$\begin{aligned} \left. \frac{\partial u}{\partial \theta} \right|_{\theta=0} &= -\frac{2R\rho}{1+\rho^2} \left(\frac{\pi}{2} + \arctan \frac{2\rho}{\rho^2-1} \right) \underset{\rho \rightarrow 0}{\approx} -R \left(\pi\rho - 4\rho^2 - \pi\rho^3 + \frac{16}{3}\rho^4 + \dots \right), \\ \left. \frac{\partial u}{\partial \theta} \right|_{\theta=\pi} &= \frac{2R\rho}{1+\rho^2} \left(\frac{\pi}{2} + \arctan \frac{2\rho}{\rho^2-1} \right) \underset{\rho \rightarrow 0}{\approx} R \left(\pi\rho - 4\rho^2 - \pi\rho^3 + \frac{16}{3}\rho^4 + \dots \right). \end{aligned}$$

The coefficients of expansion (4.16) at the point $(R, 0)$ are as follows:

$$\begin{aligned} A_0(\theta) = A_2(\theta) = 0, \quad A_1(\theta) &= \frac{4R}{\pi} \cos 2\theta, \\ A_3(\theta) &= \frac{1}{3\pi} [2R(k^2 R^2 - 4) \cos 4\theta - 4k^2 R^3 \cos 2\theta], \end{aligned} \tag{B.13}$$

$$\begin{aligned} B_0(\theta) &= C_0^{(B)} \cos \theta - \pi R \sin \theta, \quad B_1(\theta) = C_1^{(B)} \cos 2\theta - \frac{2R}{\pi} [(2\theta - \pi) \sin 2\theta + \cos 2\theta], \\ B_2(\theta) &= -\frac{k^2 R^2}{2} \cos \theta C_0^{(B)} + C_2^{(B)} \cos 3\theta - \frac{\pi R}{6} [(k^2 R^2 - 2) \sin 3\theta - 3k^2 R^2 \sin \theta], \\ B_3(\theta) &= -\frac{k^2 R^2}{3} (\sin 4\theta - 2 \sin 2\theta) C_0^{(B)} - \frac{k^2 R^2}{3} \cos 2\theta C_1^{(B)} + \cos 4\theta C_3^{(B)} \\ &\quad - \frac{4R}{3\pi} \left[\left(\frac{1}{8} k^2 R^2 - \frac{1}{2} \right) \cos^2 2\theta + \frac{1}{2} \left\{ k^2 R^2 \left(\theta - \frac{\pi}{2} \right) + 2\pi - 4\theta \right\} \sin 4\theta \right. \\ &\quad \left. - \frac{k^2 R^2}{2} \left(\pi^2 + \frac{5}{3} \right) \cos 2\theta - k^2 R^2 \left(\theta - \frac{\pi}{2} \right) \sin 2\theta + \frac{1}{4} + \frac{3k^2 R^2}{8} \left(\pi^2 - \frac{1}{6} \right) \right]. \end{aligned} \tag{B.14}$$

The asymptotic expansion at $(-R, 0)$ is obtained by an odd reflection about the y axis.

Test 6 This final test employs the boundary data with a discontinuity in the second derivative:

$$\left. \frac{\partial u}{\partial r} \right|_{r=R} = \begin{cases} \cos \varphi, & 0 < \varphi < \pi, \\ \cos 3\varphi, & \pi < \varphi < 2\pi. \end{cases}$$

In terms of (ρ, θ) , the boundary conditions read:

$$\begin{aligned} \left. \frac{\partial u}{\partial \theta} \right|_{\theta=0} &= -\frac{2R\rho}{1+\rho^2} \cos \left(\arctan \frac{2\rho}{\rho^2-1} \right) \underset{\rho \rightarrow 0}{\approx} R (-2\rho + 6\rho^3 - 10\rho^5 + \dots), \\ \left. \frac{\partial u}{\partial \theta} \right|_{\theta=\pi} &= \frac{2R\rho}{1+\rho^2} \cos \left(3 \arctan \frac{2\rho}{\rho^2-1} \right) \underset{\rho \rightarrow 0}{\approx} R (2\rho - 38\rho^3 + 170\rho^5 - \dots). \end{aligned}$$

The coefficients of expansion (4.16) at the point $(R, 0)$ are given by

$$A_0(\theta) = A_1(\theta) = A_3(\theta) = 0, \quad A_2(\theta) = \frac{32R}{3\pi} \cos 3\theta, \quad (\text{B.15})$$

$$\begin{aligned} B_0(\theta) &= C_0^{(B)} \cos \theta - 2R \sin \theta, \quad B_1(\theta) = C_1^{(B)} \cos 2\theta, \\ B_2(\theta) &= -\frac{k^2 R^2}{2} \cos \theta C_0^{(B)} + C_2^{(B)} \cos 3\theta - \frac{4R}{3\pi} \left[\{ \pi (k^2 R^2 - 6) - 32\theta \} \cos^2 \theta \sin \theta \right. \\ &\quad \left. + \left(\frac{3\pi}{2} - \pi k^2 R^2 + 8\theta \right) \sin \theta - \frac{16}{3} \cos^3 \theta + 4 \cos \theta \right], \\ B_3(\theta) &= -\frac{k^2 R^2}{3} (\sin 4\theta - 2 \sin 2\theta) C_0^{(B)} - \frac{k^2 R^2}{3} \cos 2\theta C_1^{(B)} \\ &\quad + \cos 4\theta C_3^{(B)} + \frac{k^2 R^3}{3} (4 \cos 2\theta - 3). \end{aligned} \quad (\text{B.16})$$

The asymptotic expansion at $(-R, 0)$ is obtained by an odd reflection about the y axis.